GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 15:09:37 ; Search time 358 Seconds
                 (without alignments)
                 11006.366 Million cell updates/sec

Title: US-10-621-363-12
Perfect score: 867
Sequence: 1 tcgagtttttttttttttt.........accaatctcgtgccgaattc 867

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Post-processing: Minimum Match 0%
                 Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   Published Applications NA:*
                 Listing first 45 summaries

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*	
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*	
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*	
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*	
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*	
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*	
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*	
8:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*	
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*	
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*	
11:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*	
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*	
13:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*	
14:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*	
15:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*	
16:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*	
17:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*	
18:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*	
19:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*	

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	865.2	99.8	867	10	US-09-774-639-61	Sequence 61, Appl
2	865.2	99.8	867	10	US-09-969-730-12	Sequence 12, Appl
3	865.2	99.8	867	17	US-10-621-363-12	Sequence 12, Appl
4	133.6	15.4	39703	15	US-10-017-161-1041	Sequence 1041, Ap
5	133.6	15.4	39703	16	US-10-292-798-883	Sequence 883, App
6	133.6	15.4	134292	13	US-10-240-425-1102	Sequence 1102, Ap
7	131	15.1	97247	13	US-10-087-192-688	Sequence 688, App
8	130.4	15.0	22255	9	US-09-976-740-51	Sequence 51, Appl
9	130.4	15.0	22255	13	US-10-671-242-51	Sequence 51, Appl
10	130.4	15.0	22255	14	US-10-023-529-51	Sequence 51, Appl
11	130.4	15.0	22255	14	US-10-023-523-51	Sequence 51, Appl
12	130.4	15.0	22255	16	US-10-616-187-51	Sequence 51, Appl
13	129.4	14.9	20951	16	US-09-805-455-3	Sequence 3, Appli
14	129.4	14.9	20951	17	US-10-678-140-3	Sequence 3, Appli
15	128.8	14.9	827	13	US-10-027-632-165977	Sequence 165977,
16	128.8	14.9	827	16	US-10-027-632-165977	Sequence 165977,
17	128	14.8	136726	16	US-10-085-117-244	Sequence 244, App
18	127.6	14.7	53742	13	US-10-027-192-1534	Sequence 1534, Ap
19	126.6	14.6	650	13	US-10-027-632-244049	Sequence 244049,
20	126.6	14.6	650	16	US-10-027-632-244049	Sequence 244049,
21	126.6	14.6	304	10	US-09-764-891-8441	Sequence 8441, Ap
22	126.2	14.6	149671	16	US-10-087-192-1966	Sequence 1966, Ap
23	126	14.5	727705	13	US-10-236-031B-53	Sequence 53, Appl
24	125.8	14.5	380	13	US-09-867-701-9681	Sequence 9681, Ap
25	125.8	14.5	591	16	US-10-027-632-260696	Sequence 260696,
26	125.8	14.5	591	13	US-10-027-632-260696	Sequence 260696,
27	125.8	14.5	611	16	US-10-027-632-224564	Sequence 224564,
28	125.8	14.5	611	13	US-10-027-632-224564	Sequence 224564,
29	125.6	14.5	5409	16	US-09-764-860-752	Sequence 752, App
30	125.6	14.5	5409	15	US-10-074-095-752	Sequence 752, App
31	125.6	14.5	5409	16	US-10-212-872-752	Sequence 752, App
32	125.4	14.5	585	13	US-10-027-632-189223	Sequence 189223,
33	125.4	14.5	585	16	US-10-027-632-189223	Sequence 189223,
34	125.4	14.5	30420	9	US-09-764-877-2432	Sequence 2432, Ap
35	125.4	14.5	30420	10	US-09-764-891-9164	Sequence 9164, Ap
36	125.4	14.5	30420	10	US-09-091-572-740	Sequence 740, App
37	125.4	14.5	63720	13	US-10-242-515-2432	Sequence 2432, Ap
38	125.4	14.5	63720	16	US-10-105-637-44	Sequence 44, Appl
39	125.4	14.5	63720	16	US-10-034-650-46	Sequence 46, Appl
40	125.4	14.5	96597	12	US-10-052-482-226	Sequence 226, App
41	125.4	14.5	1430	10	US-09-764-891-10178	Sequence 10178, A
42	125.4	14.5	2005	16	US-10-108-260A-2225	Sequence 2225, Ap
43	125.2	14.4	3312	13	US-10-027-632-115117	Sequence 115117,
44	125.2	14.4	3312	16	US-10-027-632-115117	Sequence 115117,
45	125.2	14.4	39776	15	US-10-160-293-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-774-639-61
Sequence 61, Application US/09774639
Publication No. US2003000355A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 90 Human Secreted Proteins
FILE REFERENCE: PZ013P1
CURRENT APPLICATION NUMBER: US/09/774,639
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US/09/774,639
PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/244,112
PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
NUMBER OF SEQ ID NOS: 371
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 61
  LENGTH: 867
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: SITE
  LOCATION: (831)
  OTHER INFORMATION: n equals a,t,g, or c
US-09-774-639-61

Query Match   99.8%; Score 865.2; DB 10; Length 867;
Best Local Similarity   100.0%; Pred. No. 6.9e-272;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGCATGAGGGTTTCACCGGTGTTAGGCCAGGAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGCATGAGGGTTTCACCGGTGTTAGGCCAGGAT 60

Qy   61 GGTCTCGGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCCTCCCAAAGTGCTGGGATTAC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGTCTCGGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCCTCCCAAAGTGCTGGGATTAC 120

Qy   121 AGGCATGAGCCACTGCGCCCAGCCGGTCTTTTTAAACATTCCCAGGACTGTACAGCCAA 180

RESULT 2
US-09-969-730-12

; Sequence 12, Application US/09969730
; Publication No. US20030054443A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; CURRENT APPLICATION NUMBER: US/09/969,730
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 09/774,639
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/238,291
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: PCT/US98/16235
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/056,371
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,732

; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,366
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,364
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,367
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,365
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,731
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,557
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,563
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/055,970
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,986
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,311
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,310
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,808
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,806
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,809
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,803
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,798
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,807
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,309
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,312
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,386
; PRIOR FILING DATE: 1997-08-05

; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 12
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (831)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-730-12

Query Match   99.8%; Score 865.2; DB 10; Length 867;
Best Local Similarity 100.0%;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAATGGGGGTTCACCGTGTTAGCCAGGAT   60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAATGGGGGTTCACCGTGTTAGCCAGGAT   60

Qy   61 GGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTAC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTAC 120

Qy   121 AGGCATGAGCCACTGCGCCCAGCCGGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGGCATGAGCCACTGCGCCCAGCCGGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAA 180

RESULT 3
US-10-621-363-12

Sequence 12, Application US/10621363
Publication No. US20040022283A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 90 Human Secreted Proteins
FILE REFERENCE: PZ013P2C1
CURRENT APPLICATION NUMBER: US/10/621,363
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/969,730
PRIOR FILING DATE: 2001-10-06
PRIOR APPLICATION NUMBER: 09/774,639
PRIOR FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: 60/238,291
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 09/244,112
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: PCT/US98/16235
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/056,371
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,732
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,366
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,364
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,370
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,373
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 373
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 12
LENGTH: 867
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (831)
OTHER INFORMATION: n equals a,t,g, or c
US-10-621-363-12

Query Match   99.8%; Score 865.2; DB 17; Length 867;
Best Local Similarity 100.0%; Pred. No. 6.9e-272;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db  721 CCAGGGCCGACTGGCCAGAGAGAGATCCGCAGAGGGCTCTGCAGCGAGCTCTGGGTGGCCAA 780
Qy  781 GCCACTCGGATTTGAACCCCGCGCTCCTCAAGTCTCAGCTGTGTTAGCCTTGANTGAAYCACC 840
Db  781 GCCACTCGGATTTGAACCCCGGGCTCCTCAAGGTCAGCTGTGTGTTAGCCTTGANTGAAYCACC 840
Qy  841 TGCTATGACCAATCTCGTGCCGAATTC 867
Db  841 TCCTATGACCAAICTCGTGCCGAATTC 867
```

RESULT 4
US-10-017-161-1041
; Sequence 1041, Application US/10017161
; Publication No. US2003014366BA1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1041
; LENGTH: 39703
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(39703)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(524)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3968)..(4102)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4517)..(4687)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5787)..(5888)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6700)..(6891)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14506)..(14748)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15377)..(15440)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16590)..(17067)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18612)..(18762)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19258)..(19390)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20773)..(20889)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20987)..(21167)
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (222295)..(22422)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23003)..(23131)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23354)..(23497)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24275)..(24444)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (393365)..(395003)
US-10-017-161-1041

Query Match   15.4%;   Score 133.6; DB 15; Length 39703;
Best Local Similarity 93.9%; Pred. No. 6.1e-32;
Matches 139; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy     6 TTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGGAGATGGGGTTCAACCGTGTTGTTAGCCAGGATGGTCT 65
Db 10987 TTTTTTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGGTTTCACCGTGTTGTTAGCCAGGATGGTCT 11046
Qy    66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCA 125
Db 11047 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCGCCTCCCAAAATGCTGGGATTACAGGCG 11106
Qy   126 TGAGCCACTGCGGCCCAGCCGGGTCTTTTTTT 153
Db 11107 TGAGCCACCGCGGCCCAGCCTTTTTTTTTTT 11134
```

RESULT 5
US-10-292-798-883
; Sequence 883, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 883
; LENGTH: 39703
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(39703)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(524)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3968)..(4102)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4517)..(4687)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5787)..(5888)
; FEATURE:
; NAME/KEY: CDS

```
;  LOCATION:        (6700)..(6891)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (14506)..(14748)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (15377)..(15440)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (16990)..(17067)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (18612)..(18762)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (19258)..(19390)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (20773)..(20889)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (20987)..(21167)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (22295)..(22422)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (24275)..(24444)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (23003)..(23131)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (23354)..(23497)
;  FEATURE:
;  NAME/KEY:        CDS
;  LOCATION:        (39365)..(39503)
```

RESULT 6
US-10-240-425-1102/c
Sequence 1102, Application US/10240425
Publication No. US20040033502A1

GENERAL INFORMATION:

```
;  APPLICANT:  Williams, Amanda
;  APPLICANT:  Boland, Joseph F.
;  APPLICANT:  Lord, Reginald V.
;  APPLICANT:  Alvarez, Chris
;  APPLICANT:  Wetzel, Jon C.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Vockley, Joseph G.
;  TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
;  FILE REFERENCE: 4492-5026
;  CURRENT APPLICATION NUMBER: US/10/240,425
;  CURRENT FILING DATE:  2002-09-30
;  PRIOR APPLICATION NUMBER: PCT/US01/09847
```

Query Match   15.4%;   Score 133.6; DB 16;   Length 39703;
Best Local Similarity   93.9%; Pred. No. 6.1e-32;
Matches 139; Conservative   0; Mismatches   9;   Indels   0;   Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTGTTAGCCAGGATGGTCT   65
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10987 TTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGACGGGGTTTCACGGTGTTGTTAGCCAGGATGGTCT   11046

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA   125
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11047 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAATGCTGGGATTACAGGCG   11106

Qy   126 TGAGCCACTGCGCCCAGCCGCGGTCTCTTTTTT   153
               |||||||||||||||||||||||||||||||||
Db   11107 TGAGCCACCGCGCCCAGCCCTTTTTTTTTT   11134

```
;  PRIOR FILING DATE: 2001-03-28
;  PRIOR APPLICATION NUMBER: US 60/193,446
;  PRIOR FILING DATE: 2000-03-31
;  NUMBER OF SEQ ID NOS: 1588
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 1102
;  LENGTH: 134292
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  OTHER INFORMATION: Genbank Accession No. US20040033502A1 AL031983
US-10-240-425-1102
```

Query Match   15.4%;   Score 133.6; DB 13;   Length 134292;
Best Local Similarity   93.9%; Pred. No. 1.1e-31;
Matches 139; Conservative   0; Mismatches   9;   Indels   0;   Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTGTTAGCCAGGATGGTCT   65
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113399 TTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGACGGGGTTTCACGGTGTTGTTAGCCAGGATGGTCT   113340

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA   125
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113339 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAATGCTGGGATTACAGGCG   113280

Qy   126 TGAGCCACTGCGCCCAGCCGCGGTCTCTTTTTT   153
               |||||||||||||||||||||||||||||||||
Db   113279 TGAGCCACCGCGCCCAGCCCTTTTTTTTTT   113252

RESULT 7
US-10-087-192-688/c
Sequence 688, Application US/10087192
Publication No. US20020182586A1

GENERAL INFORMATION:

```
;  APPLICANT:  Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 5294520001122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 688
;  LENGTH: 97247
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY:  misc_feature
;  LOCATION: (1)..(97247)
;  OTHER INFORMATION: n = A,T,C or G
US-10-087-192-688
```

Query Match   15.1%;   Score 131;   DB 13;   Length 97247;
Best Local Similarity   87.7%; Pred. No. 6.5e-31;
Matches 143; Conservative   0; Mismatches 20;   Indels   0;   Gaps   0;

Qy   2 CGAGTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTGTTAGCCAGGATG   61
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   33989 CGGCTAATTTTTTGTATTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATG   33930

Qy   62 GTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACA   121
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   33929 GTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAATGCTGGGATTACA   33870

Qy   122 GGCATGAGCCACTGCGCCCAGCCGCGGTCTCTTTTTTAAACATTCCCC   164
               || ||||||||||||||||||||||||||||||||| |||||||||
Db   33869 GGCGTGAGCCACCGCGCCCAGCCCTTTCTTAACTTCTCTCC   33827
```

RESULT 8
US-09-976-740-51
Sequence 51, Application US/09976740
Publication No. US20020194633A1
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/976,740
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 09/616,289
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 51
LENGTH: 22255
TYPE: DNA
ORGANISM: Homo sapiens
US-09-976-740-51

Query Match         15.0%;    Score 130.4;  DB 9;   Length 22255;
Best Local Similarity   89.7%;    Pred. No. 5.1e-31;
Matches  140;  Conservative   0;  Mismatches  16;  Indels   0;  Gaps   0;

Qy      6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
            ||||||||||||||||||||| |||||||| |||||||||||||||||||||||||||||
Db   7346 TTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 7405

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
            |||| ||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db   7406 CGATCTCTTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCT 7465

Qy    126 TGAGCCACTGCGCCCCAGCCGGTCTCTTTTAAACATTC 161
            |||||||| |||| |||||| ||||       ||||
Db   7466 TGAGCCACCACGCCCTGGCCTATTATTATTATTATTC 7501

---

RESULT 9
US-10-671-242-51
Sequence 51, Application US/10671242
Publication No. US20040040049A1
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/10/671,242
CURRENT FILING DATE: 2003-09-24
PRIOR APPLICATION NUMBER: US/09/616,289
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1997-11-27
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27

PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 51
TYPE: DNA
ORGANISM: Homo sapiens
US-10-671-242-51

Query Match         15.0%;    Score 130.4;  DB 13;  Length 22255;
Best Local Similarity   89.7%;    Pred. No. 5.1e-31;
Matches  140;  Conservative   0;  Mismatches  16;  Indels   0;  Gaps   0;

Qy      6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
            ||||||||||||||||||||| |||||||| |||||||||||||||||||||||||||||
Db   7346 TTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 7405

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
            |||| ||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db   7406 CGATCTCTTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCT 7465

Qy    126 TGAGCCACTGCGCCCCAGCCGGTCTCTTTTAAACATTC 161
            |||||||| |||| |||||| ||||       ||||
Db   7466 TGAGCCACCACGCCCTGGCCTATTATTATTATTATTC 7501

---

RESULT 10
US-10-023-529-51
Sequence 51, Application US/10023529
Publication No. US20020129388A1
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
APPLICANT: Lees, Robert S.
APPLICANT: Law, Simon W.
APPLICANT: Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
TITLE OF INVENTION: ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/10/023,529
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 09/616,289
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 51
LENGTH: 22255
TYPE: DNA
ORGANISM: Homo sapiens
US-10-023-529-51

Query Match         15.0%;    Score 130.4;  DB 14;  Length 22255;
Best Local Similarity   89.7%;    Pred. No. 5.1e-31;
Matches  140;  Conservative   0;  Mismatches  16;  Indels   0;  Gaps   0;

Qy      6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
            ||||||||||||||||||||| |||||||| |||||||||||||||||||||||||||||
Db   7346 TTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 7405

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
            |||| ||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db   7406 CGATCTCTTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCT 7465

Qy    126 TGAGCCACTGCGCCCCAGCCGGTCTCTTTTAAACATTC 161

RESULT 11
US-10-023-523-51
; Sequence 51, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-523-51

Query Match             15.0%;    Score 130.4;  DB 14;  Length 22255;
Best Local Similarity   89.7%;    Pred. No. 5.1e-31;
Matches  140;  Conservative   0;  Mismatches  16;   Indels    0;  Gaps   0;

Qy           6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
               |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        7346 TTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 7405

Qy          66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7406 CGATCTCTTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCT 7465

Qy         126 TGAGGCCACTGCGCCCAGCCGGGTCTCTTTTTAAACATTC 161
               |||||||||||| |||| ||| ||||  ||||||||||
Db        7466 TGAGCCACCACGCCTGGCCTATTTATTTTATTTATTC 7501

RESULT 12
US-10-616-187-51
; Sequence 51, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02

PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-616-187-51

Query Match             15.0%;    Score 130.4;  DB 16;  Length 22255;
Best Local Similarity   89.7%;    Pred. No. 5.1e-31;
Matches  140;  Conservative   0;  Mismatches  16;   Indels    0;  Gaps   0;

Qy           6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
               |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        7346 TTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 7405

Qy          66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7406 CGATCTCTTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCT 7465

Qy         126 TGAGGCCACTGCGCCCAGCCGGGTCTCTTTTTAAACATTC 161
               |||||||||||| |||| ||| ||||  ||||||||||
Db        7466 TGAGCCACCACGCCTGGCCTATTTATTTTATTTATTC 7501

RESULT 13
US-09-805-455-3/c
; Sequence 3, Application US/09805455
; Publication No. US20030166203A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001165
; CURRENT APPLICATION NUMBER: US/09/805,455
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 20951
; TYPE: DNA
; ORGANISM: Human
US-09-805-455-3

Query Match             15.0%;    Score 130.4;  DB 10;  Length 20951;
Best Local Similarity   89.7%;    Pred. No. 1.1e-30;
Matches  140;  Conservative   0;  Mismatches  16;   Indels    0;  Gaps   0;

Qy           6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
               |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        7346 TTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 7405

Qy          66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7406 CGATCTCTTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCT 7465

Qy         126 TGAGGCCACTGCGCCCAGCCGGGTCTCTTTTTAAACATTC 161
               |||||||||||| |||| ||| ||||  ||||||||||
Db        7466 TGAGCCACCACGCCTGGCCTATTTATTTTATTTATTC 7501

RESULT 14
US-10-678-140-3/c
; Sequence 3, Application US/10678140
; Publication No. US20040082772A1
; GENERAL INFORMATION:

```
;  APPLICANT: GAN, Weiniu
;  TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
;  TITLE OF INVENTION: PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL00115SDIV
;  CURRENT APPLICATION NUMBER: US/10/678,140
;  CURRENT FILING DATE: 2003-10-06
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
;
;  SEQ ID NO 3
;  LENGTH: 20951
;  TYPE: DNA
;  ORGANISM: Human
US-10-678-140-3

       Query Match             14.9%;  Score 129.4;  DB 17;  Length 20951;
       Best Local Similarity   87.1%;  Pred. No. 1.1e-30;
       Matches  142;  Conservative   0;  Mismatches  21;  Indels   0;  Gaps   0;

Qy           7 TTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTC 66
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       17276 TTTTTTTTTTTTTGTATTTTTAGTAAAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTC 17217

Qy          67 GAACTCCTGACCTCGTGATCCGCCCGCCTCGGCGCCCAAAGTGCTGGGATTACAGGCAT 126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       17216 GATCTCCTGACCTCGTGATCCGCCCGCCTCGGCGCCCAAAGTGCTGGGATTACAGGCGT 17157

Qy         127 GAGCCACTGCGCCCAGCCGGTCTCTTTAAACATTCCCCAGGAC 169
               ||||||||||||||||||||||||||||||||| |||||
Db       17156 GAGCCACCGCGCCCGGCCCGCCGTCCCTTCTTTATTATCCCACCAC 17114
```

```
RESULT 15
US-10-027-632-165977/c
;  Sequence 165977, Application US/10027632
;  Publication No. US20020198371A1
;  GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION: Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE: 2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
;
;  SEQ ID NO 165977
;  LENGTH: 827
;  TYPE: DNA
;  ORGANISM: Human
US-10-027-632-165977

       Query Match             14.9%;  Score 128.8;  DB 13;  Length 827;
       Best Local Similarity   91.9%;  Pred. No. 3.7e-31;
       Matches  136;  Conservative   0;  Mismatches  12;  Indels   0;  Gaps   0;

Qy           6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
               |||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||
Db         653 TTTTTTTTTTTTTTTGTATTTTTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 594
```

```
Qy          66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCCGGCCTCCCAAAAGTGCTGGGATTACAGGCA 125
               ||||||||||||||||| ||||||||||||||||||||| |||||||||||||||||||
Db         593 CGATCTCCTGACCTTGTGATCCGCCGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGCCG 534

Qy         126 TGAGCCACTGCGCCCAGCCGGTCTTTTTT 153
               ||||||||||||||||||||| ||||||||
Db         533 TGAACCACTGCGCCCAGCCGGTCTTTTTT 506
```

Search completed: May 24, 2004, 16:51:23
Job time : 361 secs

OM nucleic - nucleic search, using sw model

           GenCore version 5.1.6
     Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 24, 2004, 16:45:13 ; Search time 71 Seconds
                               (without alignments)
                     6776.654 Million cell updates/sec

Title: US-10-621-363-12
Perfect score: 867
Sequence: 1 tcgagtttttttttttttttt.........accaatctcgtgccgaattc 867

Scoring table: OLIGO_NUC
              Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size :   0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :   Issued_Patents_NA:*
           1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
           2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
           3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
           4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
           5: /cgn2_6/ptodata/2/ina/pCTUS_COMB.seq:*
           6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match Length	DB ID	ID	Description
1	66	7.6 148567	4	US-09-801-876B-3	Sequence 3, Appli
2	66	7.6 148567	4	US-10-254-869-3	Sequence 3, Appli
3	64	7.4 460	4	US-09-918-686-26	Sequence 26, Appl
4	64	7.4 819	3	US-09-337-171-13	Sequence 13, Appl
5	64	7.4 819	4	US-09-773-459-13	Sequence 13, Appl
6	64	7.4 92139	4	US-09-918-686-1	Sequence 1, Appli
7	61	7.0 19806	4	US-09-740-028A-3	Sequence 3, Appli
8	60	6.9 20966	4	US-09-776-976-7	Sequence 7, Appli
9	60	6.9 20966	4	US-09-909-547-7	Sequence 7, Appli
10	60	6.9 20966	4	US-09-569-852B-1	Sequence 1, Appli
11	58	6.7 506	4	US-09-621-976-15373	Sequence 15373, A
12	58	6.7 92139	4	US-09-918-686-1	Sequence 1, Appli
13	58	6.7 162450	4	US-09-345-882-1	Sequence 1, Appli
14	57	6.6 495	4	US-09-621-976-11454	Sequence 11454, A
15	57	6.6 20598	4	US-09-593-995-10	Sequence 10, Appl
16	56	6.5 13158	2	US-08-687-080-105	Sequence 105, Appl
17	56	6.5 203	4	US-09-621-976-10774	Sequence 10774, App
18	55	6.3 260	4	US-09-621-976-10668	Sequence 10668,
19	55	6.3 288	1	US-08-157-171-8	Sequence 8, Appli
20	55	6.3 289	4	US-09-621-976-10116	Sequence 10116, A
21	55	6.3 293	4	US-09-621-976-10356	Sequence 10356, A
22	55	6.3 293	4	US-09-621-976-10404	Sequence 10404, A
23	55	6.3 293	4	US-09-621-976-11002	Sequence 11002, A
24	55	6.3 293	4	US-09-621-976-11159	Sequence 11159, A
25	55	6.3 293	4	US-09-621-976-13512	Sequence 13512, A
26	55	6.3 302	4	US-09-621-976-12531	Sequence 12531, A
27	55	6.3 303	4	US-09-621-976-9840	Sequence 9840, Ap
28	55	6.3 306	4	US-09-621-976-11059	Sequence 11059, A
29	55	6.3 320	1	US-08-629-939-5	Sequence 5, Appli
30	55	6.3 320	1	US-08-759-873-5	Sequence 5, Appli
31	55	6.3 476	4	US-09-621-976-17689	Sequence 17689, A
32	55	6.3 480	4	US-09-622-976-10615	Sequence 10615, A
33	55	6.3 489	4	US-09-621-976-652	Sequence 652, App
34	55	6.3 522	4	US-09-621-976-8994	Sequence 8994, Ap
35	55	6.3 856	4	US-09-288-143-47	Sequence 47, Appl
36	55	6.3 856	4	US-09-369-247-42	Sequence 42, Appl
37	55	6.3 1268	4	US-09-157-171-3	Sequence 3, Appl
38	55	6.3 1856	3	US-08-050-159-128	Sequence 128, Appl
39	55	6.3 1856	4	US-09-702-705-321	Sequence 321, App
40	55	6.3 2280	4	US-09-614-124B-321	Sequence 321, App
41	55	6.3 2280	4	US-09-736-457-321	Sequence 321, App
42	55	6.3 2280	4	US-09-672-325-321	Sequence 321, App
43	55	6.3 2280	4	US-09-589-184-321	Sequence 321, App
44	55	6.3 3742	1	US-08-594-915-5	Sequence 5, Appli
45	55	6.3 8220	4	US-09-797-908-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-801-876B-3
; Sequence 3, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001160
; CURRENT APPLICATION NUMBER: US/09/801,876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-09-801-876B-3

Query Match   7.6%;   Score 66;   DB 4;   Length 148567;
Best Local Similarity   100.0%;   Pred. No. 1.2e-20;
Matches   66;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      70 CTCCTGACCTCGTGATCAGATCCGGCCCCCTCCGGCCTCCCAAAGTGCTGGGGATTACAGGCCATGAG 129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81514 CTCCTGACCTCGTGATCCGGCCCCCTCCGGCCTCCCAAAGTGCTGGGGATTACAGGCCATGAG 81573
```

RESULT 2
US-10-254-869-3
; Sequence 3, Application US/10254869
; Patent No. 6653117
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL00116DIV
; CURRENT APPLICATION NUMBER: US/10/254,869
; CURRENT FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 8

```
Qy     130 CCACTG 135
              ||||||
Db   81574 CCACTG 81579
```

```
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3
  LENGTH: 148567
  TYPE: DNA
  ORGANISM: Human
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (1)..(148567)
  OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3

Query Match     7.6%;  Score 66;  DB 4;  Length 148567;
Best Local Similarity   100.0%;  Pred. No. 1.2e-20;
Matches  66;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      70 CTCCTGACCTCGTGATCCGCCCGCCCGCCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCATGAG 129
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81514 CTCCTGACCTCGTGATCCGCCCGCCCGCCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCATGAG 81573

Qy     130 CCACTG 135
           ||||||
Db   81574 CCACTG 81579
```

```
RESULT 3
US-09-918-686-26
Sequence 26, Application US/0991686
Patent No. 6475739
GENERAL INFORMATION:
  APPLICANT: Brunkow, Mary
  APPLICANT: Proll, Sean
  APPLICANT: Paeper, Bryan
  APPLICANT: Staehling-Hampton, Karen
  TITLE OF INVENTION: METHODS FOR IDENTIFYING
  TITLE OF INVENTION: GENOMIC DELETIONS
  FILE REFERENCE: 240083.515
  CURRENT APPLICATION NUMBER: US/09/918,686
  CURRENT FILING DATE: 2001-07-30
  NUMBER OF SEQ ID NOS: 105
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 26
  LENGTH: 460
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-918-686-26

Query Match     7.4%;  Score 64;  DB 4;  Length 460;
Best Local Similarity   100.0%;  Pred. No. 1.3e-19;
Matches  64;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      70 CTCCTGACCTCGTGATCCGCCCGCCCGCCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCATGAG 129
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 CTCCTGACCTCGTGATCCGCCCGCCCGCCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCATGAG 375

Qy     130 CCAC 133
           ||||
Db     376 CCAC 379
```

```
RESULT 4
US-09-337-171-13/c
Sequence 13, Application US/09337171
Patent No. 6262249
GENERAL INFORMATION:
  APPLICANT: Kennedy, Giulia C
  TITLE OF INVENTION: PANCREATIC CANCER GENES
  FILE REFERENCE: 200130.454
  CURRENT APPLICATION NUMBER: US/09/337,171
  CURRENT FILING DATE: 1999-06-21
  NUMBER OF SEQ ID NOS: 15
  SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 13
  LENGTH: 819
  TYPE: DNA
  ORGANISM: Homo sapien
US-09-337-171-13

Query Match     7.4%;  Score 64;  DB 3;  Length 819;
Best Local Similarity   100.0%;  Pred. No. 1.3e-19;
Matches  64;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      70 CTCCTGACCTCGTGATCCGCCCGCCCGCCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCATGAG 129
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 CTCCTGACCTCGTGATCCGCCCGCCCGCCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCATGAG 710

Qy     130 CCAC 133
           ||||
Db     709 CCAC 706
```

```
RESULT 5
US-09-773-459-13/c
Sequence 13, Application US/09773459
Patent No. 6664054
GENERAL INFORMATION:
  APPLICANT: Kennedy, Giulia C
  TITLE OF INVENTION: PANCREATIC CANCER GENES
  FILE REFERENCE: 200130.454
  CURRENT APPLICATION NUMBER: US/09/773,459
  CURRENT FILING DATE: 2001-01-31
  PRIOR APPLICATION NUMBER: US 09/337,171
  PRIOR FILING DATE: 1999-06-21
  NUMBER OF SEQ ID NOS: 15
  SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 13
  LENGTH: 819
  TYPE: DNA
  ORGANISM: Homo sapien
US-09-773-459-13

Query Match     7.4%;  Score 64;  DB 4;  Length 819;
Best Local Similarity   100.0%;  Pred. No. 1.3e-19;
Matches  64;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      70 CTCCTGACCTCGTGATCCGCCCGCCCGCCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCATGAG 129
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 CTCCTGACCTCGTGATCCGCCCGCCCGCCCTCCGCCCTCCCAAAGTGCTGGGATTACAGGCATGAG 710

Qy     130 CCAC 133
           ||||
Db     709 CCAC 706
```

```
RESULT 6
US-09-918-686-1
Sequence 1, Application US/0991686
Patent No. 6475739
GENERAL INFORMATION:
  APPLICANT: Brunkow, Mary
  APPLICANT: Proll, Sean
  APPLICANT: Paeper, Bryan
  APPLICANT: Staehling-Hampton, Karen
  TITLE OF INVENTION: METHODS FOR IDENTIFYING
  TITLE OF INVENTION: GENOMIC DELETIONS
  FILE REFERENCE: 240083.515
  CURRENT APPLICATION NUMBER: US/09/918,686
  CURRENT FILING DATE: 2001-07-30
  NUMBER OF SEQ ID NOS: 105
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
  LENGTH: 92139
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 7043, 8369, 8401
```

```
;  OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

Query Match        7.4%;  Score 64;  DB 4;  Length 92139;
Best Local Similarity   100.0%;  Pred. No. 9.4e-20;
Matches  64;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      70 CTCCTGACCTCGTGATCCGGCCCGCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGCATGAG 129
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2380 CTCCTGACCTCGTGATCCGCCCGCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGCATGAG 2439

Qy     130 CCAC 133
           ||||
Db    2440 CCAC 2443


RESULT 7
US-09-740-028A-3
;  Sequence 3, Application US/09740028A
;  Patent No. 6410289
;  GENERAL INFORMATION:
;  APPLICANT: GONG, Fangcheng et al
;  TITLE OF INVENTION: ISOLATED HUMAN DEHYDROGENASES, NUCLEIC
;  TITLE OF INVENTION: ACID MOLECULES ENCODING THESE HUMAN DEHYDROGENASES, AND USES
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: CL001054
;  CURRENT APPLICATION NUMBER: US/09/740,028A
;  CURRENT FILING DATE: 2000-12-20
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 3
;  LENGTH: 19806
;  TYPE: DNA
;  ORGANISM: Human
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(19806)
;  OTHER INFORMATION: n = A,T,C or G
US-09-740-028A-3

Query Match        7.0%;  Score 61;  DB 4;  Length 19806;
Best Local Similarity   100.0%;  Pred. No. 2.3e-18;
Matches  61;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      60 TGGTCTCGAACTCCTGACCTCGTGATCCGGCCCGCCTCGGCCTTCCCAAAGTGCTGGGATTA 119
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     266 TGGTCTCGAACTCCTGACCTCGTGATCCGGCCCGCCTCGGCCTTCCCAAAGTGCTGGGATTA 325

Qy     120 C 120
           |
Db     326 C 326


RESULT 8
US-09-776-976-7/c
;  Sequence 7, Application US/09776976
;  Patent No. 6566332
;  GENERAL INFORMATION:
;  APPLICANT: Fruebis, Joachim
;  APPLICANT: Erickson, Mary Ruth
;  APPLICANT: Yen, Frances
;  APPLICANT: Bihain, Bernard
;  TITLE OF INVENTION: OBG3 Globular Head and Uses Thereof for Decreasing Body Mass
;  FILE REFERENCE: 76.USA4.REG
;  CURRENT APPLICATION NUMBER: US/09/776,976
;  CURRENT FILING DATE: 2001-02-05
;  PRIOR APPLICATION NUMBER: US 09/758,055
;  PRIOR FILING DATE: 2001-01-10
;  PRIOR APPLICATION NUMBER: US 60/176,228
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/198,087
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: US 60/299,881
```

```
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: Patent.pm
;  SEQ ID NO 7
;  LENGTH: 20966
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: 1..4811
;  OTHER INFORMATION: 5' regulatory region
;  NAME/KEY: misc_feature
;  LOCATION: 20560..20966
;  OTHER INFORMATION: 3' regulatory region
;  NAME/KEY: exon
;  LOCATION: 4812..4851
;  OTHER INFORMATION: exon 1
;  NAME/KEY: exon
;  LOCATION: 3787
;  OTHER INFORMATION: exon 2
;  NAME/KEY: exon
;  LOCATION: 15144..15365
;  NAME/KEY: exon
;  LOCATION: 16277..20559
;  OTHER INFORMATION: exon 3
;  NAME/KEY: misc_feature
;  LOCATION: 20560..20966
;  NAME/KEY: allele
;  LOCATION: 15120
;  OTHER INFORMATION: 9-12-48 : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 15196
;  OTHER INFORMATION: 9-12-48 : polymorphic base C or T
;  NAME/KEY: allele
;  LOCATION: 15427
;  OTHER INFORMATION: 9-12-124 : polymorphic base G or T
;  NAME/KEY: allele
;  LOCATION: 11118
;  OTHER INFORMATION: 99-14387-129 : polymorphic base A or C
;  NAME/KEY: allele
;  LOCATION: 11111
;  OTHER INFORMATION: 9-27-261 : polymorphic base G or C
;  NAME/KEY: allele
;  LOCATION: 15500
;  OTHER INFORMATION: 9-12-355 : polymorphic base G or T
;  NAME/KEY: allele
;  LOCATION: 3528..3545
;  OTHER INFORMATION: 9-16-189 : polymorphic base deletion of A
;  NAME/KEY: allele
;  LOCATION: 15073..15092
;  OTHER INFORMATION: 99-14387.rp complement
;  NAME/KEY: primer_bind
;  LOCATION: 11423..11442
;  NAME/KEY: allele
;  LOCATION: 15863
;  OTHER INFORMATION: 99-14405-105 : polymorphic base A or G
;  NAME/KEY: allele
;  LOCATION: 17170
;  OTHER INFORMATION: 9-12-428 : polymorphic base A or G
;  NAME/KEY: primer_bind
;  LOCATION: 3928..3946
;  OTHER INFORMATION: 9-27.pu
;  NAME/KEY: primer_bind
;  LOCATION: 10990..11008
;  OTHER INFORMATION: 99-14387.pu
;  NAME/KEY: primer_bind
;  LOCATION: 15503..15520
;  OTHER INFORMATION: 9-12.pu
;  NAME/KEY: primer_bind
;  LOCATION: 15759..15776
;  OTHER INFORMATION: 9-12.rp complement
;  NAME/KEY: primer_bind
;  LOCATION: 16191..16211
;  OTHER INFORMATION: 99-14405.pu
;  NAME/KEY: primer_bind
;  LOCATION: 16982..17001
;  OTHER INFORMATION: 99-14405.rp complement
```

; OTHER INFORMATION: 9-16.pu
; NAME/KEY: primer_bind
; LOCATION: 17384..17402
; OTHER INFORMATION: 9-16.rp complement
; NAME/KEY: misc_binding
; LOCATION: 3775..3799
; OTHER INFORMATION: 11106..11130
; NAME/KEY: misc_binding
; LOCATION: 99-14387-129.probe
; NAME/KEY: misc_binding
; LOCATION: 15108..15132
; OTHER INFORMATION: 9-12-48.probe
; NAME2/KEY: misc_binding
; LOCATION: 15184..15208
; OTHER INFORMATION: 9-12-124.probe
; NAME2/KEY: misc_binding
; LOCATION: 15415..15439
; OTHER INFORMATION: 9-12-355.probe
; NAME2/KEY: misc_binding
; LOCATION: 15488..15512
; OTHER INFORMATION: 9-12-428.probe
; NAME2/KEY: misc_binding
; LOCATION: 15851..15875
; OTHER INFORMATION: 99-14405-105.probe
; NAME/KEY: misc_binding
; LOCATION: 17158..17182
; OTHER INFORMATION: 9-16-189.probe
; NAME2/KEY: primer_bind
; LOCATION: 3768..3786
; OTHER INFORMATION: 9-27-261.mis
; NAME2/KEY: primer_bind
; LOCATION: 3788..3806
; OTHER INFORMATION: 9-27-261.mis complement
; NAME/KEY: primer_bind
; LOCATION: 11099..11117
; OTHER INFORMATION: 99-14387-129.mis
; NAME/KEY: primer_bind
; LOCATION: 11119..11137
; OTHER INFORMATION: 99-14387-129.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15101..15119
; OTHER INFORMATION: 9-12-48.mis
; NAME/KEY: primer_bind
; LOCATION: 15121..15139
; OTHER INFORMATION: 9-12-48.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15177..15195
; OTHER INFORMATION: 9-12-124.mis
; NAME/KEY: primer_bind
; LOCATION: 15197..15215
; OTHER INFORMATION: 9-12-124.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15408..15426
; OTHER INFORMATION: 9-12-355.mis
; NAME/KEY: primer_bind
; LOCATION: 15428..15446
; OTHER INFORMATION: 9-12-355.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15481..15499
; OTHER INFORMATION: 9-12-428.mis
; NAME/KEY: primer_bind
; LOCATION: 15501..15519
; OTHER INFORMATION: 9-12-428.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15844..15862
; OTHER INFORMATION: 99-14405-105.mis
; NAME/KEY: primer_bind
; LOCATION: 15864..15882
; OTHER INFORMATION: 99-14405-105.mis complement
; NAME/KEY: primer_bind
; LOCATION: 17151..17169
; OTHER INFORMATION: 9-16-189.mis

; NAME/KEY: primer_bind
; LOCATION: 17171..17189
; OTHER INFORMATION: 9-16-189.mis complement
US-09-776-976-7

Qy   81 GTGATCCGGCCCCCCTCCGGCCTTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCC 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10046 GTGATCCGGCCCCCCTCCGGCCTTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCC 9987

RESULT 9
US-09-909-547-7/c
; Sequence 7, Application US/0909547
; Patent No. 6579852
; GENERAL INFORMATION:
; APPLICANT: Fruebis, Joachim
; APPLICANT: Erickson, Mary Ruth
; APPLICANT: Yen, Frances
; APPLICANT: Bihain, Bernard
; TITLE OF INVENTION: OBG3 Globular Head and Uses Thereof for Decreasing Body Mass
; FILE REFERENCE: 76.US6.CIP
; CURRENT APPLICATION NUMBER: US/09/909,547
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 09/776,976
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/758,055
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: US 60/299,881
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 60/198,087
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: US 60/176,228
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent.pm
;
; SEQ ID NO 7
; LENGTH: 20966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..4811
; OTHER INFORMATION: 5' regulatory region
; NAME/KEY: exon
; LOCATION: 4812..4851
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 15144..15365
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 16277..20559
; OTHER INFORMATION: exon 3
; NAME/KEY: misc feature
; LOCATION: 20560..20966
; OTHER INFORMATION: 3' regulatory region
; NAME/KEY: allele
; LOCATION: 3787
; OTHER INFORMATION: 9-27-261 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 11118
; OTHER INFORMATION: 99-14387-129 : polymorphic base A or C
; NAME/KEY: allele
; LOCATION: 15120
; OTHER INFORMATION: 9-12-48 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 15196
; OTHER INFORMATION: 9-12-124 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 15427

; NAME/KEY: primer_bind
; LOCATION: 17171..17189
; OTHER INFORMATION: 9-16-189.mis complement
US-09-776-976-7

Query Match   6.9%; Score 60; DB 4; Length 20966;
Best Local Similarity 100.0%; Pred. No. 6.5e-18;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OTHER INFORMATION: 9-12-355 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 15500
OTHER INFORMATION: 9-12-428 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 15663
OTHER INFORMATION: 99-14405-105 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 17170
OTHER INFORMATION: 9-16-189 : polymorphic base deletion of A
NAME/KEY: primer_bind
LOCATION: 3528..3545
OTHER INFORMATION: 9-27.pu
NAME/KEY: primer_bind
LOCATION: 3928..3946
OTHER INFORMATION: 9-27.rp complement
NAME/KEY: primer_bind
LOCATION: 1090..11008
OTHER INFORMATION: 99-14387.pu
NAME/KEY: primer_bind
LOCATION: 1123..11442
OTHER INFORMATION: 99-14387.rp complement
NAME/KEY: primer_bind
LOCATION: 15073..15092
OTHER INFORMATION: 9-12.pu
NAME/KEY: primer_bind
LOCATION: 15503..15520
OTHER INFORMATION: 9-12.rp complement
NAME/KEY: primer_bind
LOCATION: 15759..15776
OTHER INFORMATION: 99-14405.pu
NAME/KEY: primer_bind
LOCATION: 1691..16211
OTHER INFORMATION: 99-14405.rp complement
NAME/KEY: primer_bind
LOCATION: 1682..17001
OTHER INFORMATION: 9-16.pu
NAME/KEY: primer_bind
LOCATION: 17384..17402
OTHER INFORMATION: 9-16.rp complement
NAME/KEY: misc_binding
LOCATION: 3775..3799
OTHER INFORMATION: 9-27-261.probe
NAME/KEY: misc_binding
LOCATION: 15184..15208
OTHER INFORMATION: 9-12-124.probe
NAME/KEY: misc_binding
LOCATION: 15415..15439
OTHER INFORMATION: 9-12-355.probe
NAME/KEY: misc_binding
LOCATION: 15488..15512
OTHER INFORMATION: 9-12-428.probe
NAME/KEY: misc_binding
LOCATION: 15108..15132
OTHER INFORMATION: 9-16-189.probe
NAME/KEY: misc_binding
LOCATION: 15851..15875
OTHER INFORMATION: 99-14405-105.probe
NAME/KEY: misc_binding
LOCATION: 17158..17182
OTHER INFORMATION: 9-16-189.probe
NAME/KEY: primer_bind
LOCATION: 3768..3786
OTHER INFORMATION: 9-27-261.mis
NAME/KEY: primer_bind
LOCATION: 3788..3806
OTHER INFORMATION: 9-27-261.mis complement
NAME/KEY: primer_bind
LOCATION: 11099..11117
OTHER INFORMATION: 99-14387-129.mis

NAME/KEY: primer_bind
LOCATION: 11119..11137
OTHER INFORMATION: 99-14387-129.mis complement
NAME/KEY: primer_bind
LOCATION: 15101..15119
OTHER INFORMATION: 9-12-48.mis
NAME/KEY: primer_bind
LOCATION: 15121..15139
OTHER INFORMATION: 9-12-48.mis complement
NAME/KEY: primer_bind
LOCATION: 15177..15195
OTHER INFORMATION: 9-12-124.mis
NAME/KEY: primer_bind
LOCATION: 15197..15215
OTHER INFORMATION: 9-12-124.mis complement
NAME/KEY: primer_bind
LOCATION: 15408..15426
OTHER INFORMATION: 9-12-355.mis
NAME/KEY: primer_bind
LOCATION: 15428..15446
OTHER INFORMATION: 9-12-355.mis complement
NAME/KEY: primer_bind
LOCATION: 15481..15499
OTHER INFORMATION: 9-12-428.mis
NAME/KEY: primer_bind
LOCATION: 15501..15519
OTHER INFORMATION: 9-12-428.mis complement
NAME/KEY: primer_bind
LOCATION: 15844..15862
OTHER INFORMATION: 99-14405-105.mis
NAME/KEY: primer_bind
LOCATION: 15864..15882
OTHER INFORMATION: 99-14405-105.mis complement
NAME/KEY: primer_bind
LOCATION: 17151..17169
OTHER INFORMATION: 9-16-189.mis
NAME/KEY: primer_bind
LOCATION: 17171..17189
OTHER INFORMATION: 9-16-189.mis complement
US-09-509-547-7

Query Match   6.98;   Score 60;   DB 4;   Length 20966;
Best Local Similarity 100.0%; Pred. No. 6.5e-18;
Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   81 GTGATCCGGCCGCCGCCTCGGCCTCCCAAAGTGCTGGGGATTACAGGCATGAGCCACTGCGGCC 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10046 GTGATCCGGCCGCCGCCTCGGCCTCCCAAAGTGCTGGGGATTACAGGCATGAGCCACTGCGGCC 9987

SEQ ID NO 1
LENGTH: 20966
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(4811)
NAME3/KEY: primer bind
LOCATION: primer bind 5' regulatory region
OTHER INFORMATION: 5' regulatory region
NAME/KEY: primer bind
LOCATION: (14683)..(14701)
OTHER INFORMATION: 17-34-860.mis
NAME/KEY: primer bind
LOCATION: (14703)..(14721)
OTHER INFORMATION: 17-34-860.mis complement
NAME/KEY: primer bind
LOCATION: (14738)..(14756)
OTHER INFORMATION: 17-34-915.mis
NAME/KEY: primer bind
LOCATION: (14758)..(14776)
OTHER INFORMATION: 17-34-915.mis complement
NAME/KEY: primer bind
LOCATION: (14796)..(14814)
OTHER INFORMATION: 17-35-306.mis
NAME/KEY: primer bind
LOCATION: (14816)..(14834)
OTHER INFORMATION: 17-35-306.mis complement
NAME/KEY: primer bind
LOCATION: (15031)..(15049)
OTHER INFORMATION: 17-35-71.mis
NAME/KEY: primer bind
LOCATION: (15051)..(15069)
OTHER INFORMATION: 17-35-71.mis complement
NAME/KEY: primer bind
LOCATION: (15101)..(15119)
OTHER INFORMATION: 9-12-48.mis
NAME/KEY: primer bind
LOCATION: (15121)..(15139)
OTHER INFORMATION: 9-12-48.mis complement
NAME/KEY: primer bind
LOCATION: (15177)..(15195)
OTHER INFORMATION: 9-12-124.mis
NAME/KEY: primer bind
LOCATION: (15197)..(15215)
OTHER INFORMATION: 9-12-124.mis complement
NAME/KEY: primer bind
LOCATION: (15408)..(15426)
OTHER INFORMATION: 9-12-355.mis
NAME/KEY: primer bind
LOCATION: (15428)..(15446)
OTHER INFORMATION: 9-12-355.mis complement
NAME/KEY: primer bind
LOCATION: (15481)..(15499)
OTHER INFORMATION: 9-12-428.mis
NAME/KEY: primer bind
LOCATION: (15501)..(15519)
OTHER INFORMATION: 9-12-428.mis complement
NAME/KEY: primer bind
LOCATION: (15661)..(15679)
OTHER INFORMATION: 17-36-47.mis
NAME/KEY: primer bind
LOCATION: (15681)..(15699)
OTHER INFORMATION: 17-36-47.mis complement
NAME/KEY: primer bind
LOCATION: (15771)..(15789)
OTHER INFORMATION: 17-36-120.mis
NAME/KEY: primer bind
LOCATION: (15791)..(15809)
OTHER INFORMATION: 17-36-120.mis complement
NAME/KEY: primer bind
LOCATION: (15844)..(15862)
OTHER INFORMATION: 99-14405-105.mis
NAME/KEY: primer bind
LOCATION: (15864)..(15882)

OTHER INFORMATION: 99-14405-105.mis complement
NAME/KEY: primer bind
LOCATION: (17151)..(17169)
OTHER INFORMATION: 9-16-189.mis
NAME/KEY: primer bind
LOCATION: (17171)..(17189)
OTHER INFORMATION: 9-16-189.mis complement
NAME/KEY: primer bind
LOCATION: (17810)..(17828)
OTHER INFORMATION: 17-37-629.mis
NAME/KEY: primer bind
LOCATION: (17830)..(17848)
OTHER INFORMATION: 17-37-629.mis complement
NAME/KEY: primer bind
LOCATION: (17992)..(18010)
OTHER INFORMATION: 17-37-811.mis
NAME/KEY: primer bind
LOCATION: (18012)..(18030)
OTHER INFORMATION: 17-37-811.mis complement
NAME/KEY: primer bind
LOCATION: (18470)..(18488)
OTHER INFORMATION: 17-38-349.mis
NAME/KEY: primer bind
LOCATION: (18490)..(18508)
OTHER INFORMATION: 17-38-349.mis complement
NAME/KEY: primer bind
LOCATION: (926)..(944)
OTHER INFORMATION: 17-30-216.mis
NAME/KEY: primer bind
LOCATION: (946)..(964)
OTHER INFORMATION: 17-30-216.mis complement
NAME/KEY: primer bind
LOCATION: (3719)..(3737)
OTHER INFORMATION: 9-27-211.mis
NAME/KEY: primer bind
LOCATION: (3739)..(3757)
OTHER INFORMATION: 9-27-211.mis complement
NAME/KEY: primer bind
LOCATION: (3754)..(3772)
OTHER INFORMATION: 9-27-246.mis
NAME/KEY: primer bind
LOCATION: (3774)..(3792)
OTHER INFORMATION: 9-27-246.mis complement
NAME/KEY: primer bind
LOCATION: (3768)..(3786)
OTHER INFORMATION: 9-27-261.mis
NAME/KEY: primer bind
LOCATION: (3788)..(3806)
OTHER INFORMATION: 9-27-261.mis complement
NAME/KEY: primer bind
LOCATION: (5076)..(5094)
OTHER INFORMATION: 17-31-298.mis
NAME/KEY: primer bind
LOCATION: (5096)..(5114)
OTHER INFORMATION: 17-31-298.mis complement
NAME/KEY: primer bind
LOCATION: (5191)..(5209)
OTHER INFORMATION: 17-31-413.mis
NAME/KEY: primer bind
LOCATION: (5211)..(5229)
OTHER INFORMATION: 17-31-413.mis complement
NAME/KEY: primer bind
LOCATION: (5364)..(5385)
OTHER INFORMATION: 17-31.rp complement
NAME/KEY: primer bind
LOCATION: (10618)..(10636)
OTHER INFORMATION: 17-32-24.mis
NAME/KEY: primer bind
LOCATION: (10638)..(10656)
OTHER INFORMATION: 17-32-24.mis complement
NAME/KEY: primer bind
LOCATION: (11020)..(11038)
OTHER INFORMATION: 99-14387-50.mis

```
;  NAME/KEY: primer_bind
;  LOCATION: (11040)..(11058)
;  OTHER INFORMATION: 99-14387-50.mis complement
;  NAME/KEY: primer_bind
;  LOCATION: (11099)..(11117)
;  OTHER INFORMATION: 99-14387-129.mis
;  NAME/KEY: primer_bind
;  LOCATION: (11119)..(11137)
;  OTHER INFORMATION: 99-14387-129.mis complement
;  NAME/KEY: primer_bind
;  LOCATION: (11169)..(11187)
;  OTHER INFORMATION: 99-14387-199.mis
;  NAME/KEY: primer_bind
;  LOCATION: (11189)..(11207)
;  OTHER INFORMATION: 99-14387-199.mis complement
;  NAME/KEY: primer_bind
;  LOCATION: (13954)..(13972)
;  OTHER INFORMATION: 17-33-TGAGACT.mis
;  NAME/KEY: primer_bind
;  LOCATION: (15144)..(15365)
;  OTHER INFORMATION: 17-33-TGAGACT.mis complement
;  NAME/KEY: primer_bind
;  LOCATION: (13974)..(13992)
;  OTHER INFORMATION: 17-33-TGAGACT.mis
;  NAME/KEY: exon
;  LOCATION: (16277)..(20559)
;  OTHER INFORMATION:
;  NAME/KEY: exon
;  LOCATION: (4812)..(4851)
;  OTHER INFORMATION:
;  NAME/KEY: misc_feature
;  LOCATION: (20560)..(20966)
;  OTHER INFORMATION: 3' regulatory region
```

```
Qy           81 GTGATCCCCCGGCTTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCC 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        10046 GTGATCCGGCCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCC 9987
```

Query Match   6.9%; Score 60; DB 4; Length 20966;
Best Local Similarity   100.0%; Pred. No. 6.5e-18;
Matches 60; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 11
US-09-621-976-15373
Sequence 15373, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15373
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15373

Query Match   6.7%; Score 58; DB 4; Length 506;
Best Local Similarity   100.0%; Pred. No. 6.6e-17;
Matches 58; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy           76 ACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCAC 133
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           51 ACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCAC 108
```

RESULT 12

US-09-918-686-1/c
Sequence 1, Application US/09918686
; Patent No. 6475739
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Proll, Sean
; APPLICANT: Paeper, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 92139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7043__8369, 8401
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

Query Match   6.7%; Score 58; DB 4; Length 92139;
Best Local Similarity   100.0%; Pred. No. 4.8e-17;
Matches 58; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy           81 GTGATCCGCCCGCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCGC 138
                |||||||||| |||||||| |||||||||||||||||||||||||||||||||||||
Db        75299 GTGATCCGCCCGGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCC 75242
```

RESULT 13
US-09-345-882-1
Sequence 1, Application US/00945882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273   : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261   : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60   : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele
```

LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 :
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-357 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-136-174 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-348 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-143-84 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-101 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146545
OTHER INFORMATION: 5-145-24 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 160031
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30

OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 93590..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58

```
;   FEATURE:
;       NAME/KEY: allele
;       LOCATION: 108084..108130
;       OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
;   FEATURE:
;       NAME/KEY: allele
;       LOCATION: 108084..108130
;       OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
;   FEATURE:
;       NAME/KEY: allele
;       LOCATION: 108084..108130
;       OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
;   FEATURE:
;       NAME/KEY: allele
;       LOCATION: 108127..108177
;       OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
;   FEATURE:
;       NAME/KEY: allele
;       LOCATION: 108127..108177
;       OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
;   FEATURE:

Query Match             6.7%;  Score 58;  DB 4;  Length 162450;
Best Local Similarity   100.0%;  Pred. No. 4.6e-17;
Matches   58;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       82 TGATCCGCCCGCCTCCGGCGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTCGCGCC 139
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   153473 TGATCCGCCCGCCTCCGGCGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTCGCGCC 153530
```

```
RESULT 14
US-09-621-976-11454/c
;   Sequence 11454, Application US/09621976
;   Patent No. 6639063
;   GENERAL INFORMATION:
;   APPLICANT: Dumas Milne Edwards, J.B.
;   APPLICANT: Jobert, S.
;   APPLICANT: Giordano, J.Y.
;   TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;   FILE REFERENCE: GENSET.054PR2
;   CURRENT APPLICATION NUMBER: US/09/621,976
;   CURRENT FILING DATE: 2000-07-21
;   NUMBER OF SEQ ID NOS: 19335
;   SOFTWARE: Patent.pm
;   SEQ ID NO 11454
;       LENGTH: 495
;       TYPE: DNA
;       ORGANISM: Homo sapiens
;   FEATURE:
;       NAME/KEY: misc_feature
;       LOCATION: 489
;       OTHER INFORMATION: n=a, g, c or t
US-09-621-976-11454

Query Match             6.6%;  Score 57;  DB 4;  Length 495;
Best Local Similarity   100.0%;  Pred. No. 1.9e-16;
Matches   57;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       68 AACTCCCTGACCTCGTCGATCCGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGGCAT 124
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 AACTCCCTGACCTCGTCGATCCGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGC 38
```

```
RESULT 15
US-09-593-995-10/c
;   Sequence 10, Application US/09593995
;   Patent No. 6406888
;   GENERAL INFORMATION:
;   APPLICANT: Conklin, Darrell C.
;   APPLICANT: Gao, Zeren
;   TITLE OF INVENTION: HELICAL CYTOKINE ZALPHA33
;   FILE REFERENCE: 99-38
;   CURRENT APPLICATION NUMBER: US/09/593,995
;   CURRENT FILING DATE: 2000-06-14
;   PRIOR APPLICATION NUMBER: 60/139,121
;   PRIOR FILING DATE: 1999-06-14
```

```
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 10
;       LENGTH: 20598
;       TYPE: DNA
;       ORGANISM: Homo sapiens
US-09-593-995-10

Query Match             6.6%;  Score 57;  DB 4;  Length 20598;
Best Local Similarity   100.0%;  Pred. No. 1.5e-16;
Matches   57;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       70 CTCCTGACCTCGTGATCCGGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCAT 126
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      242 CTCCTGACCTCGTGATCCGGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCAT 186
```

Search completed: May 24, 2004, 18:20:09
Job time : 73 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 24, 2004, 14:16:21 ; Search time 71 Seconds
[without alignments]
6776.654 Million cell updates/sec

Title: US-10-621-363-12
Perfect score: 867
Sequence: 1 tcgagtttttttttttttttt..........accaatctcgtgccgaattc 867

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0 ....accaatctcgtgccgaattc 867

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match	Length	DB	ID	Description
o	1	130.4	15.0	22255	4	US-09-616-289-51	Sequence 51, Appl
o	2	124.6	14.4	49312	4	US-09-671-317-485	Sequence 485, App
o	3	124.6	14.4	63000	4	US-09-780-172-18	Sequence 18, Appl
o	4	124.2	14.3	168575	4	US-09-426-290-1	Sequence 1, Appl
o	5	123.8	14.3	1712	3	US-09-058-389A-12	Sequence 12, Appl
o	6	123.8	14.3	1712	4	US-09-611-781-12	Sequence 12, Appl
o	7	123.8	14.3	6354	4	US-09-058-389A-5	Sequence 5, Appl
o	8	123.8	14.3	6354	4	US-09-611-781-5	Sequence 5, Appl
c	9	123.6	14.3	16063	4	US-09-801-052-3	Sequence 3, Appl
c	10	123.6	14.3	16063	4	US-10-020-121-3	Sequence 3, Appl
c	11	123	14.2	4922	2	US-08-330-272-5	Sequence 5, Appl
c	12	123	14.2	4922	5	PCT-US95-13663-5	Sequence 5, Appl
c	13	123	14.2	128779	4	PCT-US95-07855A-38	Sequence 38, Appl
o	14	122.4	14.1	174493	4	US-09-497-855A-3	Sequence 3, Appli
o	15	122.4	14.1	174493	4	US-09-004-471A-3	Sequence 3, Appli
o	16	122	14.1	152331	4	US-10-238-709-3	Sequence 3, Appli
o	17	122	14.1	176373	3	US-09-128-155-17	Sequence 17, Appl
o	18	121.8	14.0	43950	4	US-09-735-934A-3	Sequence 3, Appli
o	19	121.8	14.0	43950	4	US-10-060-332-3	Sequence 3, Appli
o	20	121.8	14.0	148567	4	US-09-801-876B-3	Sequence 3, Appli
o	21	121.8	14.0	148567	4	US-10-254-869-3	Sequence 3, Appli
o	22	121.8	14.0	429	4	US-09-621-976-9533	Sequence 9533, Ap
o	23	121.6	14.0	534	1	US-08-599-252-101	Sequence 101, App
o	24	121.6	14.0	534	5	PCT-US96-06352-101	Sequence 101, App
o	25	121.6	14.0	534	5	PCT-US96-05583-101	Sequence 101, App
o	26	121.2	14.0	8224	2	US-09-010-398-14	Sequence 14, Appl
o	27	121.2	14.0	8224	3	US-09-366-260-14	Sequence 14, Appl
c	28	121	14.0	521	4	US-09-621-976-11392	Sequence 11392, A
	29	121	14.0	1744	4	US-09-511-625B-19	Sequence 19, Appl
	30	121	14.0	17425	4	US-09-511-625B-5	Sequence 5, Appli
	31	120.8	13.9	2280	4	US-09-702-705-321	Sequence 321, App
	32	120.8	13.9	2280	4	US-09-736-457-321	Sequence 321, App
	33	120.8	13.9	2280	4	US-09-614-124B-321	Sequence 321, App
	34	120.8	13.9	2280	4	US-09-671-325-321	Sequence 321, App
	35	120.8	13.9	41100	4	US-09-589-184-321	Sequence 321, App
	36	120.8	13.9	7210	2	US-09-755-665-46	Sequence 46, Appl
c	37	120.8	13.9	7210	2	US-08-257-963B-10	Sequence 10, Appl
	38	120.6	13.9	7210	5	US-08-367-841A-10	Sequence 10, Appl
	39	120.6	13.9	14581	4	US-08-520-373D-4	Sequence 4, Appli
	40	120.6	13.9	22481	4	US-08-367-841A-43	Sequence 43, Appl
	41	120.6	13.9	22481	5	PCT-US95-07201-10	Sequence 10, Appl
	42	120.6	13.9	22481	5	PCT-US95-07201-43	Sequence 43, Appl
c	43	120.6	13.9	22484	5	US-09-875-223-2	Sequence 2, Appli
	44	120.6	13.9	22484	4	US-09-875-114-2	Sequence 2, Appli
	45	120.6	13.9	66804	4	US-09-740-041-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-616-289-51
; Sequence 51, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-51

Query Match   15.0%;   Score 130.4;   DB 4;   Length 22255;
Best Local Similarity   89.7%;
Matches 140; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy    6 TTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
          ||||||||||||||||||||  ||||||||| |||||||||||||||||||||||||||||
Db 7346 TTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 7405

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7406 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCT 7465

Qy  126 TGAGCCACTGCGCCCAGCCCGGTCTCTTTTAAACATTC 161
          ||||||||| ||||||||| |||| ||||||||||||
Db 7466 TGAGCCACCACGCCCTGGCCTCCATTTATTTATTTATTC 7501
```

RESULT 2
US-09-671-317-485

```
Sequence 485, Application US/09671317
Patent No. 6528260
GENERAL INFORMATION:
  APPLICANT: Blumenfeld, Marta
  APPLICANT: Chumakov, Ilya
  APPLICANT: Bougueleret, Lydie
  APPLICANT: Cohen, Annick
  TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
  FILE REFERENCE: 62.US3.CIP
  CURRENT FILING DATE: 2000-09-27
  CURRENT APPLICATION NUMBER: US/09/671,317
  PRIOR APPLICATION NUMBER: US 09/536,178
  PRIOR FILING DATE: 2000-03-23
  PRIOR APPLICATION NUMBER: 2000-03-24
  PRIOR FILING DATE: 2000-03-24
  PRIOR APPLICATION NUMBER: PCT/IB00/00403
  PRIOR FILING DATE: 2000-03-25
  PRIOR APPLICATION NUMBER: US 60/126,269
  PRIOR FILING DATE: 1999-03-25
  PRIOR APPLICATION NUMBER: US 60/131,961
  PRIOR FILING DATE: 1999-04-30
  NUMBER OF SEQ ID NOS: 977
  SOFTWARE: Patent.pm
SEQ ID NO 485
  LENGTH: 49312
  TYPE: DNA
  ORGANISM: Homo sapiens
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 5466..7466
  OTHER INFORMATION: 5'regulatory region
  NAME/KEY: exon
  LOCATION: 7467..7725
  OTHER INFORMATION: exon 1
  NAME/KEY: exon
  LOCATION: 20256..20355
  OTHER INFORMATION: exon 2
  NAME/KEY: exon
  LOCATION: 36905..36975
  OTHER INFORMATION: exon 3
  NAME/KEY: exon
  LOCATION: 45167..45248
  OTHER INFORMATION: exon 4
  NAME/KEY: exon
  LOCATION: 45728..45965
  OTHER INFORMATION: exon 5
  NAME/KEY: misc_feature
  LOCATION: 45966..49312
  OTHER INFORMATION: 3'regulatory region
  NAME/KEY: allele
  LOCATION: 7564
  OTHER INFORMATION: 10-286-289 : polymorphic base G or C
  NAME/KEY: allele
  LOCATION: 7619
  OTHER INFORMATION: 10-286-345 : polymorphic base A or T
  NAME/KEY: allele
  LOCATION: 7649
  OTHER INFORMATION: 10-286-375 : polymorphic base A or G
  NAME/KEY: allele
  LOCATION: 17258
  OTHER INFORMATION: 12-425-57 : polymorphic base A or G
  NAME/KEY: allele
  LOCATION: 21590
  OTHER INFORMATION: 12-421-135 : insertion of T
  NAME/KEY: allele
  LOCATION: 21595
  OTHER INFORMATION: 12-421-140 : polymorphic base A or G
  NAME/KEY: allele
  LOCATION: 36971
  OTHER INFORMATION: 12-421-140 : polymorphic base C or T
  NAME/KEY: allele
  LOCATION: 45214
  OTHER INFORMATION: 10-523-232 : polymorphic base C or T
  NAME/KEY: allele
  LOCATION: 45741
  OTHER INFORMATION: 10-289-201 : polymorphic base C or T
  OTHER INFORMATION: 10-290-37 : polymorphic base C or T
  NAME/KEY: allele
  LOCATION: 46029
  OTHER INFORMATION: 10-290-326 : polymorphic base A or G
  NAME/KEY: allele
  LOCATION: 46032
  OTHER INFORMATION: 10-290-328 : deletion of G
  NAME/KEY: primer_bind
  LOCATION: 16839..16856
  OTHER INFORMATION: 12-425.rp complement
  NAME/KEY: primer_bind
  LOCATION: 17297..17314
  OTHER INFORMATION: 12-425.rp
  NAME/KEY: primer_bind
  LOCATION: 7276..7294
  OTHER INFORMATION: 10-286.pu
  NAME/KEY: primer_bind
  LOCATION: 7676..7694
  OTHER INFORMATION: 10-286.pu complement
  NAME/KEY: primer_bind
  LOCATION: 21456..21474
  OTHER INFORMATION: 12-425.pu complement
  NAME/KEY: primer_bind
  LOCATION: 21886..21906
  OTHER INFORMATION: 12-421.pu complement
  NAME/KEY: primer_bind
  LOCATION: 36740..36758
  OTHER INFORMATION: 10-523.pu complement
  NAME/KEY: primer_bind
  LOCATION: 36997..37015
  OTHER INFORMATION: 10-523.rp complement
  NAME/KEY: primer_bind
  LOCATION: 45413..45432
  OTHER INFORMATION: 10-523.rp complement
  NAME/KEY: primer_bind
  LOCATION: 45020..45037
  OTHER INFORMATION: 10-289.pu
  NAME/KEY: primer_bind
  LOCATION: 46104..46123
  OTHER INFORMATION: 10-290.pu
  NAME/KEY: primer_bind
  LOCATION: 45705..45724
  OTHER INFORMATION: 10-290.rp complement
  NAME/KEY: primer_bind
  LOCATION: 7545..7563
  OTHER INFORMATION: 10-286-289.mis
  NAME/KEY: primer_bind
  LOCATION: 7565..7583
  OTHER INFORMATION: 10-286-289.mis complement
  NAME/KEY: primer_bind
  LOCATION: 7600..7618
  OTHER INFORMATION: 10-286-345.mis
  NAME/KEY: primer_bind
  LOCATION: 7620..7638
  OTHER INFORMATION: 10-286-345.mis complement
  NAME/KEY: primer_bind
  LOCATION: 7630..7648
  OTHER INFORMATION: 10-286-375.mis
  NAME/KEY: primer_bind
  LOCATION: 7650..7668
  OTHER INFORMATION: 10-286-375.mis complement
  NAME/KEY: primer_bind
  LOCATION: 17239..17257
  OTHER INFORMATION: 12-425-57.mis
  NAME/KEY: primer_bind
  LOCATION: 17259..17277
  OTHER INFORMATION: 12-425-57.mis complement
  NAME/KEY: primer_bind
  LOCATION: 21576..21594
  OTHER INFORMATION: 12-421-140.mis
  NAME/KEY: primer_bind
  LOCATION: 21596..21614
  OTHER INFORMATION: 12-421-140.mis complement
```

```
, NAME/KEY: primer_bind
, LOCATION: 36952..36970
, OTHER INFORMATION: 10-523-232.mis
, NAME/KEY: primer_bind
, LOCATION: 36972..36990
, OTHER INFORMATION: 10-523-232.mis complement
, NAME/KEY: primer_bind
, LOCATION: 45195..45213
, OTHER INFORMATION: 10-289-201.mis
, NAME/KEY: primer_bind
, LOCATION: 45215..45233
, OTHER INFORMATION: 10-289-201.mis complement
, NAME/KEY: primer_bind
, LOCATION: 45722..45740
, OTHER INFORMATION: 10-290-37.mis
, NAME/KEY: primer_bind
, LOCATION: 45742..45760
, OTHER INFORMATION: 10-290-37.mis complement
, NAME/KEY: primer_bind
, LOCATION: 46010..46028
, OTHER INFORMATION: 10-290-326.mis
, NAME/KEY: primer_bind
, LOCATION: 46030..46048
, OTHER INFORMATION: 10-290-326.mis complement
, NAME/KEY: misc_binding
, LOCATION: 7552..7576
, OTHER INFORMATION: 10-286-289.probe
, NAME/KEY: misc_binding
, LOCATION: 7607..7631
, OTHER INFORMATION: 10-286-345.probe
, NAME/KEY: misc_binding
, LOCATION: 7637..7661
, OTHER INFORMATION: 10-286-375.probe
, NAME/KEY: misc_binding
, LOCATION: 17246..17270
, OTHER INFORMATION: 12-425-57.probe
, NAME/KEY: misc_binding
, LOCATION: 21583..21607
, OTHER INFORMATION: 12-421-140.probe
, NAME/KEY: misc_binding
, LOCATION: 36959..36983
, OTHER INFORMATION: 10-523-232.probe
, NAME/KEY: misc_binding
, LOCATION: 45202..45226
, OTHER INFORMATION: 10-289-201.probe
, NAME/KEY: misc_binding
, LOCATION: 45729..45753
, OTHER INFORMATION: 10-290-37.probe
, NAME/KEY: misc_binding
, LOCATION: 46017..46041
, OTHER INFORMATION: 10-290-326.probe
```

```
Query Match           14.4%;     Score 124.6;  DB 4;    Length 49312;
Best Local Similarity  90.5%;     Pred. No. 7.9e-28;
Matches  133;   Conservative  0;  Mismatches  14;  Indels  0;  Gaps  0;
US-09-671-317-485

Qy          11 TTTTTTTTTTTTTTTTTAAGTAGAGAATGAGGGGTTTCACCGTGTTGTTAGCCAGGATGGTCTCGGAAC 70
               || ||||| ||| | ||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       33574 TTTTTTTATATTTTTAAGTAGAGACAGGGGTTTCACCGTGTTAGCGAGGACCGGTCT 33633

Qy          71 TCCTGACCTCGTGATCCGCCGCGGCCTCCCCAAAGTGCTGGATTACAGGCATGAGC 130
               || ||||||||||||||||||| | ||||||||||||||||||||||||||||||||
Db       33634 TCCTGACCTCGTGATTGCCCACCTCGGCCCTCCCAAAGTGCTGGGATTACAGGCATGAGC 33693

Qy         131 CACTCGCGCCCAGCCGCGGTCTTTTTAAAC 157
               |||||||||||||||||||||||| |||||
Db       33694 CACTCGCGCCCAGCCAATTATTTCAATC 33720
```

```
Patent No. 6607916
, GENERAL INFORMATION:
, APPLICANT: Robert McKay
, APPLICANT: Susan M. Freier
, APPLICANT: Jacqueline Wyatt
, TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
, FILE REFERENCE: RTS-0159
, CURRENT APPLICATION NUMBER: US/09/780,172
, CURRENT FILING DATE: 2001-02-08
, NUMBER OF SEQ ID NOS: 96
, SEQ ID NO 18
, LENGTH: 63000
, TYPE: DNA
, ORGANISM: Homo sapiens
, FEATURE:
US-09-780-172-18
```

```
Query Match           14.4%;     Score 124.6;  DB 4;    Length 63000;
Best Local Similarity  90.5%;     Pred. No. 9e-28;
Matches  133;   Conservative  0;  Mismatches  14;  Indels  0;  Gaps  0;
US-09-780-172-18

Qy           6 TTTTTTTTTTTTTTTTTTTAAGTAGAGAATGAGGGGGTTTCACCGTGTTGTTAGCCAGGATGGTCT 65
               ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       26204 TTTTTTTTTTTTTTTTTTTAAGTAGAGACAGGGGGTTTCACCGTGTTAGCCAGGACCGGTCT 26145

Qy          66 CGAACTCCTGACCTCGTGATCCGCCGCCCCGGCCTCCCCAAAGTTGCTGGGATTACAGGCA 125
               ||| ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       26144 CGATCCTCCTGACCTCGTGATCCACCCGCCCTTGGCCTCCCAAAGTGCTGGGATTACAGGTG 26085

Qy         126 TGAGCCACTGCGCCCCAGCCGCGGTCTTTT 152
               ||||||||| ||| |||||||||||| ||||
Db       26084 TGAGCCACCGCCACCCCAGCCAATTATTT 26058
```

```
RESULT 4
US-09-426-290-1/c
, Sequence 1, Application US/09426290
, Patent No. 6410712
, GENERAL INFORMATION:
, APPLICANT: Berglind Ran Olafsdottir
, APPLICANT: Jeffrey Gulcher
, TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
, FILE REFERENCE: 2345.2001-000
, CURRENT APPLICATION NUMBER: US/09/426,290
, CURRENT FILING DATE: 1999-10-25
, NUMBER OF SEQ ID NOS: 24
, SOFTWARE: FastSEQ for Windows Version 4.0
, SEQ ID NO 1
, LENGTH: 168575
, TYPE: DNA
, ORGANISM: Homo Sapiens
, FEATURE:
, NAME/KEY: CDS
, LOCATION: (21181)...(21403)
, NAME/KEY: CDS
, LOCATION: (95252)...(95430)
, NAME/KEY: CDS
, LOCATION: (101753)...(101996)
, NAME/KEY: CDS
, LOCATION: (110324)...(110439)
, NAME/KEY: CDS
, LOCATION: (124058)...(124278)
, NAME/KEY: CDS
, LOCATION: (1270009)...(127130)
, NAME/KEY: CDS
, LOCATION: (128910)...(129139)
US-09-426-290-1
```

```
Query Match           14.3%;     Score 124.2;  DB 4;    Length 168575;
Best Local Similarity  83.4%;     Pred. No. 2.1e-27;
Matches  141;   Conservative  0;  Mismatches  28;  Indels  0;  Gaps  0;
US-09-426-290-1

Qy           7 TTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGGTTCACCGTGTTGTTAGCCAGGATGGTCTC 66
```

Qy   71067 GAACTCCTGACCTCGTGATCCGGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGT 71008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 GAACTCCTGACCTCGTGATCCGGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCAT 126

Db   71007 GATCTCCTGACCTCGTGATCTCGCCCGCCTCCGGCCTCCCAAAGTGCTCGGATTACAGGCGT 70948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   127 GAGCCACTGCGCCCCAGCCGGTCCTTTTTAAACATTCCCAGGACTGTACA 175

Db   70947 GAGCCACGCGACCCCAGCCCATGTTTCTACGTTTATCCATTGATGGGCA 70899

RESULT 5
US-09-058-389A-12
; Sequence 12, Application US/0905389A
; Patent No. 6130065
; GENERAL INFORMATION:
; APPLICANT: Belt, Judith A.
; APPLICANT: Crawford, Charles R.
; APPLICANT: Patel, Divyen
; TITLE OF INVENTION: A NITROBENZYLMERCAPTOPURINERIBOSIDE   NUCLEOSIDE TRANSPORT
; TITLE OF INVENTION: (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT
; TITLE OF INVENTION: PROTEIN, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF
; TITLE OF INVENTION: USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,389A
; FILING DATE: April 9, 1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-013N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1712 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "intron 2"
; HYPOTHETICAL: NO
US-09-058-389A-12

  Query Match   14.3%;   Score 123.8;   DB 3;   Length 1712;
  Best Local Similarity   81.7%;   Pred. No. 2.3e-28;
  Matches 143;   Conservative   0;   Mismatches 32;   Indels   0;   Gaps   0;

Qy   2 CGAGTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATG 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 CGGCGGCTATTTTTTTTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATG 585

Qy   62 GTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACA 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   586 GTCTCGAATCTCCTGACCTCGTGATCCTCCCGCGTCCCGGCCTCCCAAAGTGCTGGGATTACA 645

Qy   122 GGCATGAGCCACTGCGCCCAGCCGGTCTTTTTAAACATTCCCAGGACTGTACAG 176

RESULT 6
US-09-611-781-12
; Sequence 12, Application US/09611781
; Patent No. 6423829
; GENERAL INFORMATION:
; APPLICANT: Belt, Judith A.
; APPLICANT: Crawford, Charles R.
; APPLICANT: Patel, Divyen
; TITLE OF INVENTION: A NITROBENZYLMERCAPTOPURINERIBOSIDE   NUCLEOSIDE TRANSPORT
; TITLE OF INVENTION: (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT
; TITLE OF INVENTION: PROTEIN, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF
; TITLE OF INVENTION: USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/611,781
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/058,389
; FILING DATE: April 9, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-013N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1712 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "intron 2"
; HYPOTHETICAL: NO
US-09-611-781-12

  Query Match   14.3%;   Score 123.8;   DB 4;   Length 1712;
  Best Local Similarity   81.7%;   Pred. No. 2.3e-28;
  Matches 143;   Conservative   0;   Mismatches 32;   Indels   0;   Gaps   0;

Qy   2 CGAGTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATG 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 CGGCGGCTATTTTTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATG 585

Qy   62 GTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACA 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   586 GTCTCGAATCTCCTGACCTCGTGATCCTCCCGCGTCCCGGCCTCCCAAAGTGCTGGGATTACA 645

Qy   122 GGCATGAGCCACTGCGCCCAGCCGGTCTTTTTAAACATTCAACAAACTTTTAGTGTGTCATCTACTG 176

Db   646 GGCGTGAGCCACCGTGGCCACGGCCTCATTCAACAAACTTTTAGTGTGTCATCTACTG 700

RESULT 7

US-09-058-389A-5
Sequence 5, Application US/09058389A
Patent No. 6130065
GENERAL INFORMATION:
APPLICANT: Belt, Judith A.
APPLICANT: Crawford, Charles R.
APPLICANT: Patel, Divyen
TITLE OF INVENTION: A NITROBENZYLMERCAPTOPURINERIBOSIDE
TITLE OF INVENTION: (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT
TITLE OF INVENTION: PROTEIN, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF
TITLE OF INVENTION: USE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/058,389A
FILING DATE: April 9, 1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-013N
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 6354 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
US-09-058-389A-5

Query Match 14.3%; Score 123.8; DB 3; Length 6354;
Best Local Similarity 81.7%; Pred. No. 4.6e-28;
Matches 143; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
Qy    2 CGAGTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATG 61
          | |||||||||||||||||||||   |||||||||||||||||||
Db 1111 CGCGGCGTATTTTTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATG 1170

Qy   62 GTCTCGAACTCCTGACCTCGTGATCCGCGCCGCCTCGGCCTTCCCAAAGTGCTGGGATTACA 121
          ||||| ||||||||||||||||| ||||||||||||||||||| ||||||||||||||||||
Db 1171 GTCTCGATCTCCTGACCTCGTGATCCTCCCGCGCCTCCCAAAGTGCTGGGATTACA 1230

Qy  122 GGCATGAGCCACTGCGCCCAGGCGTCTTTTTAAACATTCCCAGGACTGTACAG 176
          ||||| |||||||| |||| |||||||||||||||||| ||||||||| |||
Db 1231 GGCGTGAGCCACCGTGCCCAGCCTCATTCAACAAACTTTTAGTGTGCATCTACTG 1285
```

RESULT 8
US-09-611-781-5, Application US/09611781
Sequence 5, Application US/09611781
Patent No. 6423829
GENERAL INFORMATION:
APPLICANT: Belt, Judith A.
APPLICANT: Crawford, Charles R.
APPLICANT: Patel, Divyen
TITLE OF INVENTION: A NITROBENZYLMERCAPTOPURINERIBOSIDE
TITLE OF INVENTION: (NBMPR)-INSENSITIVE, EQUILIBRATIVE, NUCLEOSIDE TRANSPORT

TITLE OF INVENTION: PROTEIN, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF
TITLE OF INVENTION: USE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/611,781
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/058,389
FILING DATE: April 9, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-013N
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 6354 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
US-09-611-781-5

Query Match 14.3%; Score 123.8; DB 4; Length 6354;
Best Local Similarity 81.7%; Pred. No. 4.6e-28;
Matches 143; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
Qy    2 CGAGTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATG 61
          | |||||||||||||||||||||   |||||||||||||||||||
Db 1111 CGCGGCGTATTTTTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATG 1170

Qy   62 GTCTCGAACTCCTGACCTCGTGATCCGCGCCGCCTCGGCCTTCCCAAAGTGCTGGGATTACA 121
          ||||| ||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db 1171 GTCTCGATCTCCTGACCTCGTGATCCTCCCGCGCCTCCCAAAGTGCTGGGATTACA 1230

Qy  122 GGCATGAGCCACTGCGCCCAGGCGTCTTTTTAAACATTCCCAGGACTGTACAG 176
          ||||| |||||||| |||| |||||||||||||||||| ||||||||| |||
Db 1231 GGCGTGAGCCACCGTGCCCAGCCTCATTCAACAAACTTTTAGTGTGCATCTACTG 1285
```

RESULT 9
US-09-801-052-3/c, Application US/09801052
Sequence 3', Application US/09801052
Patent No. 6368842
GENERAL INFORMATION:
APPLICANT: BRASLEY, Ellen
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: CL001045
CURRENT APPLICATION NUMBER: US/09/801,052
CURRENT FILING DATE: 2001-03-08
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 16063

```
;   TYPE: DNA
;   ORGANISM: Human
US-09-801-052-3

Query Match         14.3%;  Score 123.6;  DB 4;  Length 16063;
Best Local Similarity   93.5%;  Pred. No. 8.8e-28;
Matches 129;  Conservative  0;  Mismatches  9;  Indels  0;  Gaps  0;

Qy      8 TTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCG  67
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    222 TTTTTTTTTTGTATTTTTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCG 163

Qy     68 AACTCCTGACCTCGTGATCCGCCCGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 ATCTCCTGACCTCGTGATCCGCCCGCCGTCTAGGCCCCCAAAAAGTGCTGGGATTACAGGCATG 103

Qy    128 AGCCACTGCGCCCAGCCG 145
            |||||| |||||||||
Db    102 AGCCACCCGCGCCAGCCG  85
```

RESULT 10
US-10-020-121-3/c
; Sequence 3, Application US/10020121
; Patent No. 6638747
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001045DIV
; CURRENT APPLICATION NUMBER: US/10/020,121
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/255,386
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/801,052
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-020-121-3

```
Query Match         14.3%;  Score 123.6;  DB 4;  Length 16063;
Best Local Similarity   93.5%;  Pred. No. 8.8e-28;
Matches 129;  Conservative  0;  Mismatches  9;  Indels  0;  Gaps  0;

Qy      8 TTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCG  67
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    222 TTTTTTTTTTGTATTTTTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCG 163

Qy     68 AACTCCTGACCTCGTGATCCGCCCGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 ATCTCCTGACCTCGTGATCCGCCCGCCGTCTAGGCCCCCAAAAAGTGCTGGGATTACAGGCATG 103

Qy    128 AGCCACTGCGCCCAGCCG 145
            |||||| |||||||||
Db    102 AGCCACCCGCGCCAGCCG  85
```

RESULT 11
US-08-330-272-5
; Sequence 5, Application US/08330272
; Patent No. 5985598
; GENERAL INFORMATION:
; APPLICANT: Russo et al
; TITLE OF INVENTION: TCL-1 Gene and Protein and Related
; TITLE OF INVENTION: Methods and Compositions
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds

;   STREET: 1155 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: U.S.A.
;   ZIP: 10036
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; APPLICATION NUMBER: US/08/330,272
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Misrock, S. Leslie
;   REGISTRATION NUMBER: 18,872
;   REFERENCE/DOCKET NUMBER: 6754-027
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 790-9090
;   TELEFAX: (212) 790-8864/9741
;   TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4922 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA
US-08-330-272-5

```
Query Match         14.2%;  Score 123;  DB 2;  Length 4922;
Best Local Similarity   85.2%;  Pred. No. 7.1e-28;
Matches 132;  Conservative  4;  Mismatches  19;  Indels  0;  Gaps  0;

Qy      6 TTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT  65
            ||||||||||||||||||:|||||  |||||:||  ||||||||||||||||||||||||
Db   3291 TTTTTTTTTTTWTWTWTWTTTTTAGTAGAGMCGGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 3350

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
            ||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||||
Db   3351 CGATCTCCTGACCTCATGATCTGCCCGCCCTCGGGCCCCCAAAAGTGCTGGGATTACAGGCA 3410

Qy    126 TGAGCCACTGCGCCCAGCCGCGCGGTCTTTTTAAACATT 160
            |||||||||||||||||||  |||||||||| ||||||
Db   3411 TGAGCCACCACGTCCCGGCCCTTACCATTGCTTTATT 3445
```

RESULT 12
PCT-US95-13663-5
; Sequence 5, Application PC/TUS9513663
; GENERAL INFORMATION:
; APPLICANT: Russo et al
; TITLE OF INVENTION: TCL-1 Gene and Protein and Related
; TITLE OF INVENTION: Methods and Compositions
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13663
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

```
;       NAME: Mistock, S. Leslie
;       REGISTRATION NUMBER: 18,872
;       REFERENCE/DOCKET NUMBER: 6754-027
;  TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (212) 790-9090
;       TELEFAX: (212) 790-8864/9741
;       TELEX: 66141 PENNIE
;  INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;       LENGTH: 4922 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: unknown
;       TOPOLOGY: unknown
;       MOLECULE TYPE: DNA
PCT-US95-13663-5

Query Match             14.2%;  Score 123;  DB 5;  Length 4922;
Best Local Similarity   85.2%;  Pred. No. 7.1e-28;
Matches  132; Conservative   4; Mismatches  19; Indels   0; Gaps   0;

Qy       3291 TTTTTTTTTTT|TTT|T|TTTTWTWTWWTTTTTTTTAGTAGAGAGACAGATGGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 3350
              ||||||||||:|:|||:|:||||:|||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3291 TTTTTTTTTTT|TTT|T|TTTWTWTWWTTTTTTTTAGTAGAGAGAGMCGGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 3350

Qy         66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3351 CGAACTCCTGACCTCATGATCTGCCCGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCA 3410

Qy        126 TGAGCCACTGCGCCCAGCCGGGTCTTTTTAAACATT 160
              ||||||||||||||||||||||||||||||
Db       3411 TGAGCCACCACAGTCGGCCCTTACCATTGCTTTATT 3445
```

```
RESULT 13
US-09-497-855A-38/c
;  Sequence 38, Application US/0949785SA
;  Patent No. 6605432
;  GENERAL INFORMATION:
;  APPLICANT: Huang, Tim
;  TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
;  FILE REFERENCE: UM01523
;  CURRENT APPLICATION NUMBER: US/09/497,855A
;  CURRENT FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/120,592
;  PRIOR FILING DATE: 1999-02-18
;  PRIOR APPLICATION NUMBER: 60/118,760
;  PRIOR FILING DATE: 1999-02-05
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: Patentin version 3.0
;  SEQ ID NO 38
;  LENGTH: 128779
;  TYPE: DNA
;  ORGANISM: Homo sapiens;
US-09-497-855A-38

Query Match             14.2%;  Score 123;  DB 4;  Length 128779;
Best Local Similarity   92.8%;  Pred. No. 4.2e-27;
Matches  129; Conservative   0; Mismatches  10; Indels   0; Gaps   0;

Qy          6 TTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGAGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17434 TTTTTTTTTTTTTTTGTATTTTAGTAGAGAACGGGGTTTCACCATGTTAGCCAGGATGGTCT 17375

Qy         66 CGAACTCCTGACCTCGTGATCCGGCCTCGGGCCTCCCAAAGTTGCTGGGATTACAGGCA 125
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17374 CGAACTCCTGACCTCGTGATCCGCCCGACTCGGGCCTCCCAAAGTTGCTGGGATTACAGGGC 17315

Qy        126 TGAGCCACTGCGCCCAGCC 144
              ||||||||||||||||||
Db      17314 TGAGCCACTGCCCGGCC 17296
```

```
RESULT 14
```

```
US-09-804-471A-3/c
;  Sequence 3, Application US/09804471A
;  Patent No. 6479269
;  GENERAL INFORMATION:
;  APPLICANT: WEBSTER, Marion et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: CL001164
;  CURRENT APPLICATION NUMBER: US/09/804,471A
;  CURRENT FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 3
;  LENGTH: 174493
;  TYPE: DNA
;  ORGANISM: Human
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(174493)
;  OTHER INFORMATION: n = A,T,C or G
US-09-804-471A-3

Query Match             14.1%;  Score 122.4;  DB 4;  Length 174493;
Best Local Similarity   89.2%;  Pred. No. 7.6e-27;
Matches  132; Conservative   0; Mismatches  16; Indels   0; Gaps   0;

Qy          6 TTTTTTTTTTTTTTTTTTAAGTAGAGATGAGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              ||||||||||||||||||  || ||||||||| |||||||||||||||||||||||||||||
Db      89001 TTTTTTTTTTTTTTGTATTCTAGTACAGATGGGGTTTCACCATGTTGGCCAGGATGGTCT 88942

Qy         66 CGAACTCCTGACCTCGTGATCCGGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db      88941 CGAGTCCTGACCTCATGATCTGCCCGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGGG 88882

Qy        126 TGAGGCACTGCGCCCAGCCGGGTCTTTTTT 153
              ||||||||||||||||||||||| |||||
Db      88881 TGAGCCACTGCGCCCAGCCGGTCTTTTCT 88854
```

```
RESULT 15
US-10-238-709-3/c
;  Sequence 3, Application US/10238709
;  Patent No. 6680188
;  GENERAL INFORMATION:
;  APPLICANT: WEBSTER, Marion et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: CL001164DIV
;  CURRENT APPLICATION NUMBER: US/10/238,709
;  CURRENT FILING DATE: 2002-09-11
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 3
;  LENGTH: 174493
;  TYPE: DNA
;  ORGANISM: Human
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(174493)
;  OTHER INFORMATION: n = A,T,C or G
US-10-238-709-3

Query Match             14.1%;  Score 122.4;  DB 4;  Length 174493;
Best Local Similarity   89.2%;  Pred. No. 7.6e-27;
Matches  132; Conservative   0; Mismatches  16; Indels   0; Gaps   0;

Qy          6 TTTTTTTTTTTTTTTTTTAAGTAGAGATGAGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              ||||||||||||||||||  || ||||||||| |||||||||||||||||||||||||||||
Db      89001 TTTTTTTTTTTTTTGTATTCTAGTACAGATGGGGTTTCACCATGTTGGCCAGGATGGTCT 88942

Qy         66 CGAACTCCTGACCTCGTGATCCGGCCCGCCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
```

Db   88941 CGATCTCCTGACCTCATGATCTGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGGG 88882
                ||| |||||||||||| ||||| ||||| ||||||||||||||||||||||||||||||
Qy   126 TGAGCCACTGCGCCCAGCCGGGTCTTTTT 153
                ||||||||||||||||||||||||| ||
Db   88881 TGAGCCACTGCGCCCAGCCTGTTTTTCTT 88854

Search completed: May 24, 2004, 15:50:37
Job time : 74 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 24, 2004, 14:13:41 ; Search time 2357 Seconds
(without alignments)
10984.519 Million cell updates/sec

Title: US-10-621-363-12
Perfect score: 867
Sequence: 1 tcgagtttttttttttttttt.........accaatctcgtccgaattc 867

Scoring table: IDENTIFY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   EST:*

```
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_esthum:*
18: em_estin:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_inv:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_pro:*
26: em_gss_rod:*
27: em_gss_phg:*
28: em_gss_vrl:*
29: gb_gss1:*
29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	560	64.6	586	14	CA392561	CA392561 cs26b04.y
2	560	64.6	586	14	CA392562	CA392562 cs26b05.y
3	502.6	58.0	527	14	CA866165	CA866165 ir39a06.x
4	468.4	54.0	491	10	AW193512	AW193512 xm17b12.x

c	No.	Score	%	Length	DB	ID	Description
c	5	435.8	50.3	941	13	BQ710765	BQ710765 AGENCOURT
c	6	409	47.2	421	9	AI521901	AI521901 t180d03.x
c	7	253	29.2	392	9	AI493646	AI493646 th36b01.x
c	8	133.6	15.4	574	28	AQ387243	AQ387243 RPCI11-15
c	9	131.4	15.2	615	13	BX504260	BX504260 DKFZp686E
c	10	129.8	15.0	599	28	AQ423092	AQ423092 CITB1-B1
c	11	129	14.9	699	9	BU689261	BU689261 UI-CF-ECl
c	12	128.8	14.9	292	9	AA599141	AA599141 aes2d02.s
c	13	128.8	14.8	507	14	CD171998	CD171998 UI-H-FH1-
c	14	128.8	14.8	647	13	BU619183	BU619183 UI-H-FH1-
c	15	128.8	14.8	3538	11	BC041578	BC041578 Homo sapi
c	16	128.4	14.8	378	10	AW264969	AW264969 AGENCOURT
c	17	128.4	14.8	452	13	BU838849	BU838849 AGENCOURT
c	18	128.4	14.8	457	13	BQ681574	BQ681574 AGENCOURT
c	19	128.2	14.8	641	13	BU629053	BU629053 UI-H-FL0-
c	20	128	14.8	669	29	AG150408	AG150408 Pan trogl
c	21	127.8	14.7	246	13	BX475938	BX475938 DKFZp686G
c	22	127.8	14.7	972	10	BF308728	BF308728 60189023
c	23	127	14.6	591	28	AQ634459	AQ634459 RPCI-11-4
c	24	127	14.6	715	12	BI668873	BI668873 603294840
c	25	127	14.6	759	29	AG185133	AG185133 Pan trogl
c	26	126.8	14.6	808	14	CD654994	CD654994 AGENCOURT
c	27	126.8	14.6	228	10	AW265197	AW265197 xq53e02.x
c	28	126.8	14.6	354	13	BX645228	BX645228 DKFZp781I
c	29	126.4	14.6	468	28	AQ220641	AQ220641 HS_2006_B
c	30	126.4	14.6	487	28	B52695	B52695 CIT-HSP-200
c	31	126.4	14.6	672	29	BM664570	BM664570 UI-E-CLi
c	32	126.4	14.6	684	29	AG178266	AG178266 Pan trogl
c	33	126.4	14.6	155	14	BU537276	BU537276 AGENCOURT
c	34	126.2	14.6	216	14	CB055058	CB055058 NISC_gm07
c	35	126.2	14.6	222	9	AV738722	AV738722 AV738722
c	36	126.2	14.6	298	10	BF061274	BF061274 7d55h07.x
c	37	126.2	14.6	304	10	BE676952	BE676952 7d49b07.x
c	38	126.2	14.6	417	9	AU148124	AU148124 AU148124
c	39	126.2	14.6	419	28	AQ123954	AQ123954 HS_3103_A
c	40	126.2	14.6	487	13	BX641257	BX641257 DKFZp586F
c	41	126.2	14.6	506	14	CB055059	CB055059 NISC_gm07
c	42	126.2	14.6	630	29	AG011645	AG011645 Homo sapi
c	43	126.2	14.6	720	28	AQ893423	AQ893423 HS_3148_A
c	44	126.2	14.6	1018	12	BG675124	BG675124 602621192
c	45	126	14.5	300	13	BQ436628	BQ436628 AGENCOURT

ALIGNMENTS

RESULT 1
CA392561/c

```
LOCUS       CA392561      586 bp    mRNA    linear   EST   06-NOV-2002
DEFINITION  cs26b04.y1 Human Retinal pigment epithelium/choroid cDNA
            cs26b04.y1 Homo sapiens cDNA clone cs26b04
            (Un-normalized, unamplified; cs Homo sapiens cDNA clone cs26b04
            5', mRNA sequence.
ACCESSION   CA392561
VERSION     CA392561.1  GI:24725399
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 586)
AUTHORS     Wistow,G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE       Expressed sequence tag analysis of human RPE/choroid for the
            NEIBank Project: Over 6000 non-redundant transcripts, novel genes
            and splice variants
JOURNAL     Mol. Vis. 8 (4), 205-220 (2002)
MEDLINE     22103460
PUBMED      12107410
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
```

Pax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 26 row: b column: 04
Seq primer: M13RP1 reverse primer (ABI).

FEATURES   Location/Qualifiers
    source   1..586
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="cs26b04"
                  /tissue_type="RPE/choroid"
                  /dev_stage="Adult"
                  /lab_host="EMDH10B"
                  /clone_lib="Human Retinal pigment epithelium/choroid cDNA
                  (Un-normalized, unamplified): cs"
                  /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
                  eyes (75-80 years old) yielded approximately 600 mg of
                  dissected RPE/choroid tissue. This in turn yielded 340 ug
                  of total RNA and 7 ug of mRNA. A directionally cloned cDNA
                  library in the pCMVSPORT6 vector was constructed at Life
                  Technologies (Rockville, MD; now part of Invitrogen Corp),
                  essentially following the protocols of the SuperScript
                  Plasmid System (Invitrogen Corp.
                  <http://www.invitrogen.com/>). The library code
                  designation was cs. For this library, cDNA inserts were
                  cloned into the NotI/MluI sites of the vector. EST
                  analysis was performed on the unamplified library at the
                  NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match   64.6%;   Score 560; DB 14; Length 586;
Best Local Similarity   99.3%;   Pred. No. 4e-96;
Matches 582; Conservative   1; Mismatches   1; Indels   2; Gaps   2;

Qy   27 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   586 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 527

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCCAGCCCG 146
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCCAGCCCG 467

Qy   147 TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACATTTGGGAAC 206
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466 TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACATTTGGGAAC 407

Qy   207 TCCCCCCACGGCCATAACTGATCTCCAGAGGTAAGACCAAGACAAGAATGGGGGGATTC 266
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 TCCCCCCACGGCCATAACTGATCTCCAGAGGTAAGACCAAGACAAGAATGGGGGGATTC 347

Qy   267 ACATCTAAGGTCTGGTCGATGGCTGATGAAGGAAGAGAAGAATCAGCGGAACAAAAGCCCTCTAG 326
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 ACATCTAAGGTCTGGTCGATGGCTGATGAAGGAAGAGAAGAATCAGCGGAACAAAAGCCTTCTAG 287

Qy   327 GTCTTTCTTACCACAAAACACCCTCTGCCCACCTCGCTTTGAAAGGGGCGCAGAAGTATAGTG 386
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 GTCTTTCTTACCACAAAACACCCTCTGCCCACCTCGCTTTGAAAGGGGCGCAGAAGTATAGTG 227

Qy   387 GGCGA-GCTGCCCACCTGGTCACAGTGAAGGGGATCTGGAGAAATACTCACACTTGAGGTG 445
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226 GGCGAGGCTGCCCACCTGGTCACAGTGAAGGGGATCTGGAGAAATACTCACACTTGAGGTG 167

Qy   446 CTCGCCCTCTTCATCAGCCAGTCTAACTTAAGCCAATGACCCACCGGGAG-CTTACACA 504
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   166 CTCGCCCTCTTCATCAGCCAGTCTAACTTAAGCCAATGACCCACCGGGACCCTTACACA 107

Qy   505 AGTYCAAACACAGGCCCAAAATGCATTCATTGAGCAGGAGGGGCCCAAAGGACTCCGGACGGAGA 564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 AGTCCAAACACAGGCCCAAAATGCATTCATTGAGCAGGAGGGGCCCAAAGGACTCCGGACGGAGA 47

Qy   565 GAGGGCCCAATAAGGCTGGGGCGTTATTTCCGATCCATAGAGAGAGCAG 610
          ||||||||||||||||||||||||||||||||||||||||||||||| 1
Db   46 GAGGGCCCAATAAGGCTGGGGCGTTATTTCCGATCCATAGAGAGAGCAG 1

RESULT 2

LOCUS   CA392562   586 bp   mRNA   linear   EST 06-NOV-2002
DEFINITION   CA392562/c cs26b05.Y1 Human Retinal pigment epithelium/choroid cDNA
             (Un-normalized, unamplified): cs Homo sapiens cDNA clone cs26b05
             5', mRNA sequence.
ACCESSION   CA392562
VERSION   CA392562.1 GI:24725401
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 586)
  AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,
             Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE   Expressed sequence tag analysis of human RPE/choroid for the
             NEIBank Project: Over 6000 non-redundant transcripts, novel genes
             and splice variants
  JOURNAL   Mol. Vis. 8 (4), 205-220 (2002)
  MEDLINE   22103460
   PUBMED   12107410
COMMENT   Contact: Wistow G
             Section on Molecular Structure and Function
             National Eye Institute
             6/331, NIH, Bethesda, MD 20892-2740, USA
             Tel: 301 402 3452
             Fax: 301 496 0078
             Email: graeme@helix.nih.gov
             Plate: 26 row: b column: 05
             Seq primer: M13RP1 reverse primer (ABI).

FEATURES   Location/Qualifiers
    source   1..586
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="cs26b05"
                  /tissue_type="RPE/choroid"
                  /dev_stage="Adult"
                  /lab_host="EMDH10B"
                  /clone_lib="Human Retinal pigment epithelium/choroid cDNA
                  (Un-normalized, unamplified): cs"
                  /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
                  eyes (75-80 years old) yielded approximately 600 mg of
                  dissected RPE/choroid tissue. This in turn yielded 340 ug
                  of total RNA and 7 ug of mRNA. A directionally cloned cDNA
                  library in the pCMVSPORT6 vector was constructed at Life
                  Technologies (Rockville, MD; now part of Invitrogen Corp),
                  essentially following the protocols of the SuperScript
                  Plasmid System (Invitrogen Corp.
                  <http://www.invitrogen.com/>). The library code
                  designation was cs. For this library, cDNA inserts were
                  cloned into the NotI/MluI sites of the vector. EST
                  analysis was performed on the unamplified library at the
                  NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match   64.6%;   Score 560; DB 14; Length 586;
Best Local Similarity   99.3%;   Pred. No. 4e-96;
Matches 582; Conservative   1; Mismatches   1; Indels   2; Gaps   2;

Qy   27 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   586 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 527

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCCAGCCCG 146
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCCAGCCCG 467

Qy   147 TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACATTTGGGAAC 206
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
RESULT 3
CA866165
LOCUS       CA866165                527 bp    mRNA    linear    EST 20-DEC-2002
DEFINITION  ir39a06.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6547308 3',
            mRNA sequence.
ACCESSION   CA866165
VERSION     CA866165.1  GI:27317714
KEYWORDS    EST.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 527)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lenishka,I., Sceace,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., Mccann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1..527
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6547308"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN
```

```
Query Match      58.0%;   Score 502.6;   DB 14;   Length 527;
Best Local Similarity   99.4%;   Pred. No. 3e-85;
Matches  524;   Conservative   1;   Mismatches   0;   Indels   2;   Gaps   2;

Qy     22 TTTTAAGTAGAGGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCG  81
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCG  60

Qy     82 TGATCCGGCCCGCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCA  141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGATCCGGCCCGCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCA  120

Qy    142 GCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCCAACCCATACTACCTGACACATTTG  201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCCAACCCATACTACCTGACACATTTG  180

Qy    202 GGAACTCCCCCCACGGCCCATAACTGATCTCGCAGGTAAGACCAAGAGCAAGAAATGGGG  261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGAACTCCCCCCACGGCCCATAACTGATCTCGCAGGTAAGACCAAGAGCAAGAAATGGGG  240
```

Email: cgaps-r@email.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco

```
FEATURES             Location/Qualifiers
     source          1..491
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2684447"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /clone_lib="NCI_CGAP_Ut4"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.48 Kb. Life Technologies catalog #:
                     1542-016"
ORIGIN
Query Match          54.0%;   Score 468.4;   DB 10;   Length 491;
Best Local Similarity   99.6%;   Pred. No. 9.1e-79;
Matches  480;   Conservative  0;   Mismatches  1;   Indels  1;   Gaps  1;

Qy    19  TTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCCAGGCTGGTCTCGAACTCCTGACC  78
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  TTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCCAGGCTGGTCTCGAACTCCTGACC  60

Qy    79  TCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGC  138
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCC  120

Qy   139  CCAGCCGGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACTGACAT  198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CCAGCCGGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACTGACAT  180

Qy   199  TTGGGAACTCCCCCCACGGCCCATAACTGATCTGCAGAGGTAAGCACAAGAGCAAGAAGTG  258
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  TTGGGAACTCCCCCCACGGCCCATAACTGATCTGCAGAGGTAAGCACAAGAGCAAGAAGTG  240

Qy   259  GGGGGATTCACATCTAAGGTGTTGGTAATGGTCTGATGAATGAAGAGAGACGGAACAAAA  318
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GGGGGATTCACATCTAAGGTGTTGGTAATGGTCTGATGAATGAAGAGAGACGGAACAAAA  300

Qy   319  GCCCTCTAGGGTCGTTTTCTTACCACAAACACCTCTCGCCCACCTGCTTTGAAGGGGCCAGAA  378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GCCCTCTAGGGTCTTTTCTTACCACAAACACCTCTCGCCCACCTGCTTTGAAGGGGCCAGAA  360

Qy   379  GTATAGTGGGCCGA-GCTGCCCACCTCGCTACAGTGAAGGGATCTGAGAAATACTCACACT  437
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   361  GTATAGTGGGGCGGCTGCCCACCTCGCTACAGTGAAGGGATCTGAGAAATACTCACACT  420

Qy   438  TTGAGGTGCTCCGCCCGCCTCTTCATCAGCGCCAGTCCACCAATGACCCCACGGGGAGC  497
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  TTGAGGTGCTCCGCCCGCCTCTTCATCAGCGCCAGTCCACCAATGACCCCACGGGGAGC  480

Qy   498  TT  499
          ||
Db   481  CT  482
```

RESULT 5
```
LOCUS       BQ710765    941 bp    mRNA    linear    EST  16-JUL-2002
DEFINITION  AGENCOURT_8484953 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6301113
            5', mRNA sequence.
ACCESSION   BQ710765
VERSION     BQ710765.1  GI:21849664
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 941)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgaps-r@email.nih.gov
            Tissue Procurement: Dr. Mark Watson
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2516  row: h  column: 10
            High quality sequence stop: 583.
FEATURES             Location/Qualifiers
     source          1..941
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6301113"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_113"
                     /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following: 5' adaptor:
                     GCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN
Query Match          50.3%;   Score 435.8;   DB 13;   Length 941;
Best Local Similarity   96.0%;   Pred. No. 1e-72;
Matches  454;   Conservative  2;   Mismatches  16;   Indels  1;   Gaps  1;

Qy   382  TAGTGGGGCGAGCTGCCCACCTCGCTACAGTGAAGGGATCTGAGAAATACTCACACTTTGA  441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   713  TGGGGGGGAGGGCTGCCCCCCTTGNTACAGGGAAGGGATCTGAGAAATACTCACACTTTGA  654

Qy   442  GGTGCTCGCCCCTTCATCAGCCAGCTCTTAACTTAAGCCAATGACCCCACGGGGAG-CTTA  500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   653  GGTGCTCGCCCCTTCATCAGCCAGCTCTTAACTTAAGCCAATGACCCCACGGGGAGCCTTC  594

Qy   501  CACAAGTTCAAACAGGCCCAAAATNCATTCATTCATGNCAGGGAGCCCAAAGGACTCCGGGAG  560
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   593  CACAAGTTCCAAACAGGCCCAAATGCATTCATGAGCAGGGAGGGCCCAAAGGACTCCGGAG  534

Qy   561  GAGAGAGGGCCCAATAAGGCTGGTGCCTATTTCCGGATCCATGAGAGAGGAGAGCACGGAGCAG  620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   533  GAGAGAGGGCCCAATAAGGCTGGTGCCTATTTCCGGATCCATGAGAGAGGAGAGCACGGAGCAG  474

Qy   621  GCCCTTTGATTAAGTATATCATTCTTGAATGCAAGCTTCAAAATCCGGTATGCCGGGTG  680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   473  GCCCTTTGATTAANGTATACATTCTTGAATGCAAGCTTCAAAATCCGGTATGCCGGGTG  414

Qy   681  AGAATGAGCAGGACTAACACCTGGGTGTCATGGGCAAGCCTCCAGGGCCGGACTGGGCAG  740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   413  AGAATGAGCAGGACTAACACCTGGGTGTCATGGGCAAGCCTCCAGGGCCGGACTGGGCGGAGAG  354

Qy   741  ACAGANTCCGCAAAGAGAGCTCTGCAGCCAGCTCTGGTGCCAAGCCACTCGGANTTGAACCCC  800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   353  ACAGANTCCGCAAAGAGAGCTTGCAGCCAGCTCTGGTGCCAAGCCACTCGGANTTGAACCCC  294

Qy   801  GGCTCCTCAAGGTCAGCTGTGTAGCCTTGANTGAAYCACCTGCTATGACCAAT  853
```

```
Db    293 GGCTTCCTCAAGGTCAGCTGTGTGTAGCCTTGAATGAATGAATCACCTTGCTATGACCAAT 241
```

RESULT 6
LOCUS   AI521901   421 bp   mRNA   linear   EST 13-APR-1999
DEFINITION ti80d03.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2138309 3',
            similar to contains Alu repetitive element; contains element PTR5
            ; mRNA sequence.
ACCESSION   AI521901
VERSION   AI521901.1 GI:4436036
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 421)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 401.
            Insert Length: 718   Std Error: 0.00
FEATURES   Location/Qualifiers
     source   1..421
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2138309"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIds 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo."
ORIGIN
      Query Match   47.2%; Score 409; DB 9; Length 421;
      Best Local Similarity 99.8%; Pred. No. 1.7e-67;
      Matches 420; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy     22 TTTTAAGTAGAGAATGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCG  81
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTAAGTAGAGGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCG  60

Qy     82 TGATCCGCCCGCCTCGGCCTCCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCA 141
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCA 120

Qy    142 GCCGGCTTCTTTTAAACATTCCCCAGGACTGTTACAGCCAACCCATACTCACTCGACATTG 201
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCGGTCTTTTTAAACATTCCCAGGACGTATCAGCCAACCCATACTCACTCGACATTG 180

Qy    202 GGAACTCCCCCACGGCCATAACTGATCTCGCAGAGGTAAGAGCAAGAGCAAGAATGGGG 261
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    293 GGCTTCCTCAAGGTCAGCTGTGTGTAGCCTTGAATGAATGAATCACCTTGCTATGACCAAT ...
```

RESULT 7
LOCUS   AI493546   392 bp   mRNA   linear   EST 30-MAR-1999
DEFINITION th3eb01.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2120329 3',
            similar to contains Alu repetitive element.; mRNA sequence.
ACCESSION   AI493546
VERSION   AI493546.1 GI:4394549
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 392)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11548-013
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/bbrp/image/image.html
            Seq Primer: -40UP from Gibco
            High quality sequence stop: 391.
            Insert Length: 882   Std Error: 0.00
FEATURES   Location/Qualifiers
     source   1..392
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2120329"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pan1"
                     /tissue_type="adenocarcinoma"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.72 kb. Life Technologies catalog #:
                     11548-013"
ORIGIN
      Query Match   29.2%; Score 253; DB 9; Length 392;
      Best Local Similarity 100.0%; Pred. No. 5.6e-38;
      Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      6 TTTTTTTTTTTTTTTTTTTTAAGTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT  65
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 199

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125

Qy    200 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 259
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    200 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 259
```

Db   181 GGAACTCCCCCACGGGCCCATAACTGATCTCGCAGAGGTAAGAGCAAGAGCAAGAATGGGG 240
Db   262 GATTCACATCTAAGGTTCTGGTGATGGGCTGGAAGGAAGGAAGAATCAGCGAACAAAAGCC 321
Db   241 GATTCACATCAAGGTCTGGTGATCAGCGTGATAAGGAAGGAAGAATCAGCGAACAAAAGCC 300
Qy   322 TCTAGGTCTTTCTTACCACAAACACCTTCTTCTGCCCACCTGGTTTGAAAGGGCCAGAGTA 381
Db   301 TCTAGGTCTTTCTTACCACAAACACCTTCTTCTGCCCACCTGGTTTGAAAGGGCCAGAGTA 360
Qy   382 TAGTGGGCGA-GCTGCCCACCTGCTACAGTGAAGGGATTCTGGAGAAATACTCACACTTG 440
Db   361 TAGTGGGGCGAGGCTGCCCACCTGCTACAGTGAAGGGATTCTGGAGAAATACTCACACTTG 420
Qy   441 A 441
Db   421 A 421

```
Qy    126 TGAGCCACTGCGCCCAGCCGGTCTTTTT 153
          |||||| ||||||||||||| | |||||
Db    147 TGAGCCACCCCGCCCAGCCGCTTTTTTTT 174
```

RESULT 8
```
AQ387243
LOCUS       AQ387243                574 bp    DNA     linear   GSS 21-MAY-1999
DEFINITION  RPCI11-154N12.TV RPCI-11 Homo sapiens genomic clone RPCI-11-154N12,
            genomic survey sequence.
ACCESSION   AQ387243
VERSION     AQ387243.1  GI:4358266
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSS: RPCI11-154N12.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@deejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..574
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7559075"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-154N12"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
ORIGIN
```

```
Query Match   15.4%;   Score 133.5;  DB 28;   Length 574;
Best Local Similarity   93.9%;   Pred. No. 1.8e-15;
Matches 139;   Conservative  0;   Mismatches  9;   Indels  0;   Gaps  0;
```

```
Qy      6 TTTTTTTTTTTTTTTTTTTAAGTAGAGATAGAGGGGGTTCACCGTGTTAGCCAGGATGGTCT 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27 TTTTTTTTTTTTTTTTTTTAGTAGAGAGAGAGGGGTTTCACCGTGTTAGCCAGGATGGTCT 86

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     87 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAAATGCTGGGATTACACGCG 146
```

RESULT 9
```
Qy    126 TGAGCCACTGCGCCCAGCCGGTCGTCTTT 153
          |||||| |||||||||||||| | | |||
Db    147 TGAGCCACCCCGCCCAGCCGCTTTTTTTT 174
```

```
BX504260
LOCUS       DKFZp686E20133_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION  DKFZp686E20133 5'', mRNA sequence.
            BX504260                615 bp    mRNA    linear   EST 04-SEP-2003
ACCESSION   BX504260
VERSION     BX504260.1  GI:32030985
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 615)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No si sequence available.
            This clone (DKFZp686E20133) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1..615
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686E20133"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN
```

```
Query Match   15.2%;   Score 131.4;  DB 13;   Length 615;
Best Local Similarity   99.2%;   Pred. No. 4.6e-15;
Matches 132;   Conservative  0;   Mismatches  1;   Indels  0;   Gaps  0;
```

```
Qy      2 CGAGTGTTTTTTTTTTTTTTTTTTTTAAGTAGAGATAGGGGGTTTCACCGTGTTAGCCAGGATG 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 CGATTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTTCACCGTGTTAGCCAGGATG 77

Qy     62 GTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACA 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 GTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAAGTGCTGGGATTACA 17

Qy    122 GGCATGAGCCACT 134
          ||||||||||||
Db     16 GGCATGAGCCACT 4
```

RESULT 10
```
AQ423092/c
LOCUS       AQ423092                590 bp    DNA     linear   GSS 23-MAR-1999
DEFINITION  CITB1-E1-2575M23.TR CITB1-E1 Homo sapiens genomic clone 2575M23,
            genomic survey sequence.
ACCESSION   AQ423092
```

```
VERSION     AQ423092.1  GI:4480816
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shiruya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other GSSs: CITBI-E1-2575M23.TF
            Contact: Shaying Zhao, William Nierman, Mark Adams
            The Institute for Genomic Research
            Department of Eukaryotic Genomics
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..590
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2575M23"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     Caltech Human BAC Library D"
ORIGIN
Query Match          15.0%;  Score 129.8;  DB 28;  Length 590;
Best Local Similarity 95.0%;  Pred. No. 9.5e-15;
Matches  134;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;

Qy    4 AGTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGT 63
         |||||||||||||||||||| ||||||||||||||||||||||||| ||||||||||||||
Db  364 ATTTTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCTTGTTAGCCAGGATGGT 305

Qy   64 CTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGTCGGGATTACAGG 123
         ||| |||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  304 CTTGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG 245

Qy  124 CATGAGCCACTGCGCCCAGCC 144
         |||||||||||||||||| ||
Db  244 CATGAGCCACTGCGCCCGGCC 224
```

```
RESULT 11
BU689261
LOCUS       BU689261                699 bp    mRNA    linear   EST 07-OCT-2002
DEFINITION  UI-CF-EC1-adw-o-18-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
            UI-CF-EC1-adw-o-18-0-UI 3', mRNA sequence.
ACCESSION   BU689261
VERSION     BU689261.1  GI:23546857
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97054477
```

```
PUBMED      8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            The following repetitive elements were found in this cDNA
            sequence: 7-137, >ALU (matched compliment)
            Seq primer: M13 FORWARD
            PolyA=Yes.
FEATURES             Location/Qualifiers
     source          1..699
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EC1-adw-o-18-0-UI"
                     /tissue_type="Lung"
                     /dev_stage="Adult and Fetal"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-EC1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-EC1 is a normalized cDNA library containing the
                     following tissue(s): Normal lung from adult and from fetal
                     day 64, day 87, week 19 and week 42. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     AAGTGCTTAC.
                     TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
                     and 380-383
                     TAG_LIB=UI-CF-EC1
                     TAG_SEQ=AAGTGCTTAC"
ORIGIN
Query Match          14.9%;  Score 129;  DB 13;  Length 699;
Best Local Similarity 96.4%;  Pred. No. 1.3e-14;
Matches  132;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy    8 TTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGTCTCG 67
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    1 TTTTTTTTTTTTTTTTTTTTTCAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCG 60

Qy   68 AACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
         | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 120

Qy  128 AGCCACTGCGCCCAGCC 144
         ||||||||| |||||||
Db  121 AGCCACTGCACCCAGCC 137
```

```
RESULT 12
AA599141
LOCUS       AA599141                292 bp    mRNA    linear   EST 06-MAR-1998
DEFINITION  ae52d02.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone
            IMAGE:950499 3' similar to contains Alu repetitive element; contains
            element MER40 repetitive element ;, mRNA sequence.
```

ACCESSION   AA599141
VERSION   AA599141.1   GI:2432766
KEYWORDS   EST.
SOURCE   Homo sapiens
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 292)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
             Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
             Martin,J., Moore,B., White,Y., Schellenberg,K., Steptoe,M., Tan,F.,
             Theising,B., Wylie,T., Waterston,R. and Wilson,R.
  TITLE   WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Wilson RK
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert Length: 400   Std Error: 0.00
             Seq primer: -40m13 fwd. ET from Amersham.
FEATURES   Location/Qualifiers
     source   1..292
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:950499"
                     /tissue_type="lung carcinoma"
                     /cell_line="NCI-H69"
                     /dev_stage="cell line NCI-H69"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene lung carcinoma 937218"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
                     Oligo dT, Small cell carcinoma cell line NCI-H69. Average
                     insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5'
                     CTCGAG:TTTTTTTTTTTTTTTTTT 3'."
ORIGIN

 Query Match   14.9%;   Score 128.8; DB 9;   Length 292;
 Best Local Similarity   80.3%;   Pred. No. 2e-14;
 Matches 151; Conservative   0; Mismatches 37; Indels   0; Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGAATGGGGGTTCACCGTGTTAGCCAGGATGGTCT 65
            ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1 TTTTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 60

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
            |||| ||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   1 TTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 60

Qy   61 TGATCTCCCTGACCTCGTGATCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGTG 120
            ||||||||||||||||| |||||||||| |||||||||||||||||||||||||||||||
Db   126 TGAGCCACTGCGTCCCACCCGGTTCTTTTAACATTCCCACGGACTGTACAGCCACCAT 185

Qy   121 TGAGCCACTGCGTCCCACCCGTTCTTTCAACTAATTCTTTCCTCACTATTCCAGCTGC 180

Qy   186 ACTCACCT 193
            ||||||||
Db   181 ACTTGTCT 188

RESULT 13
CD171998/c
LOCUS   CD171998   507 bp   mRNA   linear   EST 19-MAY-2003
DEFINITION AGENCOURT_14065742 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30375733 5', mRNA sequence.
ACCESSION   CD171998
VERSION   CD171998.1   GI:30853082
KEYWORDS   EST.
SOURCE   Homo sapiens (human)

ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 507)
  AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE   NIH-MGC http://mgc.nci.nih.gov/.
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM437 row: d column: 14
            High quality sequence stop: 507.
FEATURES   Location/Qualifiers
     source   1..507
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30375733"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

 Query Match   14.8%;   Score 128.6; DB 14;   Length 507;
 Best Local Similarity   93.7%;   Pred. No. 1.7e-14;
 Matches 134; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   10 TTTTTTTTTTTTTTTTTTTAAGTAGAGAGAATGGGGGTTCACCGTGTTAGCCAGGATGGTCGAA 69
            ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db   505 TTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCGAA 446

Qy   70 CTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAG 129
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||| |||
Db   445 CTCCTGACCTCGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAG 386

Qy   130 CCACTGCGCGCCCAGCCCGGTCTTTT 152
            |||||||||||||||| |||||||||
Db   385 CCACCGCGCGCCCAGCCCAAATTTTT 363

RESULT 14
BU619183
LOCUS   BU619183   647 bp   mRNA   linear   EST 23-SEP-2002
DEFINITION UI-H-FH1-bfn-k-19-0-UI.s1 NCI_CGAP_FH1 Homo sapiens cDNA clone
            UI-H-FH1-bfn-k-19-0-UI 3', mRNA sequence.
ACCESSION   BU619183
VERSION   BU619183.1   GI:22285398
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 647)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares: bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 6-136, >ALU (matched compliment)
Seq primer: M13 FORWARD

FEATURES   Location/Qualifiers
     source   1. .647
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FH1-bfn-k-19-0-UI"
                     /tissue_type="Cell Line"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FH1"
                     /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
                     (Pharmacia) with a modified polylinker; Site 1: EcoR I;
                     Site_2: Not I; NCI_CGAP_FH1 is a normalized cDNA library
                     obtained from a cell line derived from grade I
                     chondrosarcoma tissue. The library was constructed and
                     normalized according to Bonaldo, Lennon and Soares, Genome
                     Research, 6:791-806, 1996. First strand cDNA synthesis was
                     primed with an oligo-dT primer containing a Not I site.
                     Double stranded cDNA was ligated to an EcoR I adaptor,
                     digested with Not I, and cloned directionally into
                     pT7T3-Pac vector. The oligonucleotide used to prime the
                     synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     AGAATCCGGC. The cell line was provided by Dr. James Martin
                     from the University of Iowa.
                     TAG TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
                     Chondrosarcoma
                     TAG_LIB=UI-H-FH1
                     TAG_SEQ=AGAATCCGGC"
POLYA=yes.

ORIGIN

Query Match   14.8%; Score 128.6; DB 13; Length 647;
Best Local Similarity   90.7%; Pred. No. 1.5e-14;
Matches 137; Conservative   0; Mismatches 14; Indels   0; Gaps   0;

Qy   9 TTTTTTTTTTTTTTTTTTAAGTAGTAGAGAGAGGGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGA 68
        |||||||||||||||||| ||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1 TTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTTGA 60

Qy   69 ACTCCTGACCTTGTGATCCGCCCGCCTCGGCGTTCCCAAAGTGCTGGGATTACAGCCATGA 128
        |||||||||| ||||||| |||||||||| |||||||||||||||||||| ||||||| ||
Db   1 TCTCCTGACCTCGTGATCTCGCCCGCCTCCGGCCTCCCAAAGTGCTGGGACTACAGGCGTGA 120

Qy 129 GCCACTGCGCCCCAGCCGGGTCTTTTTAAACAT 159
        ||||| | ||||||||| || ||||||||| ||
Db 121 GCCACCACGCCCAGCCACCAGTTTTTAAAAT 151

RESULT 15
BC041578/c
LOCUS   BC041578   3538 bp   mRNA   linear   HTC 19-NOV-2003
DEFINITION Homo sapiens hypothetical protein MGC17986, mRNA (cDNA clone
            IMAGE:4640152), with apparent retained intron.
ACCESSION   BC041578
VERSION   BC041578.1 GI:27469657
KEYWORDS   HTC.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3538)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED   12477932
  MEDLINE   22338257
REFERENCE   2 (bases 1 to 3538)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-DEC-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,N.L., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 44 Row: e Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23957699
            This clone has the following problem: retained intron.

FEATURES   Location/Qualifiers
     source   1. .3538
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4640152"
                     /tissue_type="Skin, melanotic melanoma."
                     /clone_lib="NIH_MGC_20"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"

ORIGIN

Query Match   14.8%; Score 128.6; DB 11; Length 3538;
Best Local Similarity   93.7%; Pred. No. 7.5e-15;
Matches 134; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   9 TTTTTTTTTTTTTTTTTTTTAAGTAGTAGAGACGATGGGGGGTTTCACCGTGTTAGCCGTTAGCCAGGATGGTCTTGA 68
        |||||||||||||||||||| ||||||||||||| |||||||||||| |||||||||||||||||||||||||||||
Db 2222 TTTTTTTTTTTTTTTTTTTTTAAGTAGTAGAGACGGGGTTTCACCGTGTTATCACCGTGTTAGCCAGGATGGTCTTGA 2163

```
Qy    69 ACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGA 128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2162 TCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGA 2103

Qy   129 GCCACTGCGCCCAGCCGGTCTTT 151
         |||||| ||| ||| |||||||||
Db  2102 GCCACCGCGGGCGGGCAGGTCTTT 2080
```

Search completed: May 24, 2004, 15:49:07
Job time : 2362 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: May 24, 2004, 16:42:18 ; Search time 2339 Seconds
(without alignments)
11069.051 Million cell updates/sec

Title: US-10-621-363-12
Perfect score: 867
Sequence: 1 tcgagtttttttttttttt............accaatctcgtgccgaattc 867

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST:*

			%		Query		
	1:	em_estba:*					
	2:	em_esthum:*					
	3:	em_estin:*					
	4:	em_estmu:*					
	5:	em_estov:*					
	6:	em_estpl:*					
	7:	em_estro:*					
	8:	em_htc:*					
	9:	gb_est1:*					
	10:	gb_est2:*					
	11:	gb_htc:*					
	12:	gb_est3:*					
	13:	gb_est4:*					
	14:	gb_est5:*					
	15:	em_estfun:*					
	16:	em_estom:*					
	17:	em_gss_hum:*					
	18:	em_gss_inv:*					
	19:	em_gss_pln:*					
	20:	em_gss_vrt:*					
	21:	em_gss_fun:*					
	22:	em_gss_mam:*					
	23:	em_gss_mus:*					
	24:	em_gss_pro:*					
	25:	em_gss_rod:*					
	26:	em_gss_phg:*					
	27:	em_gss_vrl:*					
	28:	gb_gss1:*					
	29:	gb_gss2:*					

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB ID		Description
1	374	43.1	491	10	AW193512	AW193512 xm17b12.x
2	371	42.8	421	9	AI521901	AI521901 ti80d03.x
3	371	42.8	527	14	CA866165	CA866165 ir39a06.x
4	366	42.2	586	14	CA392561	CA392561 cs26b04.y

	5	366	42.2	271	13	CA392562	CA392562 cs26b05.y
c	6	253	31.3	941	13	BQ210765	BQ210765 AGENCOURT
c	7	129	29.2	392	9	AI493546	AI493546 th36b01.x
c	8	129	14.9	615	13	BX504260	BX504260 DXFZp686E3
c	9	76	8.8	507	9	AA611893	AA611893 no81h09.s
c	10	76	8.8	518	9	AV719687	AV719687
c	11	76	8.8	592	9	AV718792	AV718792
c	12	76	8.8	601	9	AV719326	AV719326
c	13	76	8.8	604	9	AV720014	AV720014
c	14	76	8.8	613	9	AV720842	AV720842
c	15	74	8.5	481	12	BM561574	BM561574 AGENCOURT
c	16	74	8.5	579	9	AU145055	AU145055 AU145055
c	17	74	8.5	693	29	AG088636	AG088636 Pan trogl
c	18	71	8.2	269	9	AI648558	AI648558 tz55f08.x
c	19	71	8.2	590	28	AQ423092	AQ423092 CITB1-E1-
c	20	71	8.2	621	29	AG161021	AG161021 Pan trogl
c	21	70	8.1	736	29	AG034583	AG034583 Pan trogl
c	22	68	7.8	334	9	AI431956	AI431956 th32g11.x
c	23	68	7.8	569	14	CA438955	CA438955 UI-H-DIO-
c	24	68	7.8	871	13	BX437247	BX437247 UI-H-DIO-
c	25	67	7.7	446	28	AQ083222	AQ083222 RPCI11-54
c	26	67	7.7	839	12	BG164166	BG164166 602s43050
c	27	66	7.6	489	9	AL119649	AL119649 DXFZp761K
c	28	65	7.5	317	13	BU532671	BU532671 AGENCOURT
c	29	65	7.5	352	13	BF832090	BF832090 PM3-HT090
c	30	65	7.5	480	14	H98660	H98660 yx17a06.s1
c	31	65	7.5	507	28	AQ169564	AQ169564 HS_3176_B
c	32	65	7.5	636	29	AG127873	AG127873 Pan trogl
c	33	65	7.5	674	28	AQ314972	AQ314972 RPCI11-94
c	34	65	7.5	878	13	BU192901	BU192901 AGENCOURT
c	35	64	7.4	245	10	BG010083	BG010083 PM4-GN030
c	36	64	7.4	285	14	T58118	T58118 yb26d09.r1
c	37	64	7.4	286	10	BF901544	BF901544 IL2-MT017
c	38	64	7.4	334	14	AI357477	AI357477 qu01f06.x
c	39	64	7.4	366	9	AV759382	AV759382 AV759382
c	40	64	7.4	374	10	AW500250	AW500250 UI-HF-BN0
c	41	64	7.4	390	28	AQ240731	AQ240731 CIT-HSP-2
c	42	64	7.4	403	28	AQ586609	AQ586609 RPCI-11-4
c	43	64	7.4	406	10	AW440424	AW440424 hb90c11.x
c	44	64	7.4	411	9	AI355224	AI355224 qt79c08.x
c	45	64	7.4	413	12	BI492842	BI492842 df30a07.w

ALIGNMENTS

RESULT 1
AW193512

LOCUS AW193512 491 bp mRNA linear EST 29-NOV-1999
DEFINITION xm17b12.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2684447 3',
 similar to contains Alu repetitive element ;, mRNA sequence.
 repetitive element /, mRNA sequence.
ACCESSION AW193512
VERSION AW193512.1 GI:6472211
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 491)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

ORIGIN
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 395.

FEATURES   Location/Qualifiers
     source   1..491
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone=IMAGE:2684447"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Ut4"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.48 kb. Life Technologies catalog #:
                     11542-016"

ORIGIN

Query Match   43.1%; Score 374; DB 10; Length 491;
Best Local Similarity 100.0%; Pred. No. 3.8e-53;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    19 TTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGTCTCGGAACTCCTGACC 78
         ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     1 TTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGTCTCGAACTCCTGACC 60

Qy    79 TCGTGATCCGCGCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGC 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TCGTGATCCGCGCCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGC 120

Qy   139 CCAGCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACAT 198
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCAGCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACAT 180

Qy   199 TTGGGAACTCCCCACCCCACGGCCCATAACTGATCTCGAGAGGTAAGACCAAGAAGAATG 258
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTGGGAACTCCCCACGGCCCATAACTGATCTCGAGAGGTAAGACCAAGAAGAATG 240

Qy   259 GGGGATTCACATCTAAGGTCTGGTGATGGCTGATCAAGGAAGAGAATCAGCGAACAAAA 318
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGGGATTCACATCTAAGGTCTGGTGATGGCTGATCAAGGAAGAGAATCAGCGGAACAAAA 300

Qy   319 GCCTCTAAGGTCTTTCTTACCACAAACACCTCTCTCGCCCACCTGCTTTGAAAGGGGCAGAA 378
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GCCTCTAAGGTCTTTCTTACCACAAACACCTCTCTCGCCCACCTGCTTTGAAAGGGGCAGAA 360

Qy   379 GTATAGTGGGCGAG 392
         ||||||||||||||
Db   361 GTATAGTGGGCGAG 374
```

RESULT 2
AI521901
LOCUS   AI521901   421 bp   mRNA   linear   EST 13-APR-1999
DEFINITION ti80d03.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2138309 3'
            similar to contains Alu repetitive element;contains element PTR5
            repetitive element ;, mRNA sequence.
ACCESSION   AI521901
VERSION   AI521901.1 GI:1436036
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 421)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmett-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 718   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 401.

FEATURES   Location/Qualifiers
     source   1..421
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2138309"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIds 1322376-1322911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "

ORIGIN

Query Match   42.8%; Score 371; DB 9; Length 421;
Best Local Similarity 100.0%; Pred. No. 1.4e-52;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    22 TTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCG 81
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCG 60

Qy    82 TGATCCGCCGCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGATGAGCCACCGCGCCCA 141
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TGATCCGCCGCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGATGAGCCACCGCGCCCA 120

Qy   142 GCCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACATTTG 201
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACATTTG 180

Qy   202 GGAACTCCCCCACGGCCCATAACTGATCTCGCAGAGGTAAGACCAAGAACAAGAGATGGGG 261
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGAACTCCCCCACGGCCCATAACTGATCTCGCAGAGGTAAGACCAAGAACAAGAGATGGGG 240

Qy   262 GATTCACATCTAAGGTCTCGGTGATGGCTGATGAGGAAGAGAAGAATCAGCGAACAAAAGCC 321
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GATTCACATCTAAGGTCTCGGTGATGGCTGATGAGGAAGAGAAGAATCAGCGAACAAAAGCC 300

Qy   322 TCTAAGGTCTTTCTTACCACAAACACCTCTCTGCCCACCTGCTTTGAAAGGGGCAGAAGTA 381
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TCTAAGGTCTTTCTTACCACAAACACCTCTCTGCCCACCTGCTTTGAAAGGGGCAGAAGTA 360

Qy   382 TAGTGGGCGAG 392
         |||||||||||
Db   361 TAGTGGGCGAG 371
```

RESULT 3
CA866165
LOCUS   CA866165   527 bp   mRNA   linear   EST 20-DEC-2002
DEFINITION ir39a06.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6547308 3',
            mRNA sequence.
ACCESSION   CA866165
VERSION   CA866165.1 GI:27317714
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 527)
AUTHORS     Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,B., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
TITLE       Endocrine Pancreas Consortium
JOURNAL     Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1..527
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6547308"
                     /tissue_type="Purified pancreatic islet"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     Not1; Site_2: Xho1; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     Xho1 site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave. St.
                     Louis, MO 63110, E-mail: hinoue@ingate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN

Query Match     42.8%;  Score 371;  DB 14;  Length 527;
Best Local Similarity   100.0%;  Pred. No. 1.1e-52;
Matches  371;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   22 TTTTAAGTAGAGGATGGGGTTTCACCGGTTGAGCGAGGATGGTCTCGGAACTCCTGACCTCG 81
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TTTTAAGTAGAGGATGGGGTTTCACCGGTTGAGCGAGGATGGTCTCGGAACTCCTGACCTCG 60

Qy   82 TGATCCGCGCCCCTCGGCCTCCCAAAGTGTGGGATTACAGGCATGAGCCACTCGCCCCA 141
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGATCCGCGCCCCTCGGCCTCCCAAAGTGTGGGATTACAGGCATGAGCCACTCGCCCCA 120

Qy  142 GCCGGTCTTTTTAAACATTTCCCAGACACTGTACAGCCAACCCATACTCACCTGACCATTTG 201
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GCCGGTCTTTTTAAACATTTCCCAGACACTGTACAGCCAACCCATACTCACCTGACCATTTG 180

Qy  202 GGAACTCCCCCCACGGCCCATAACTGATCTCCAGAGACGTAAGACCAAGACAAGAATGGGG 261
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GGAACTCCCCCCACGGCCCATAACTGATCTCCAGAGACGTAAGACCAAGACAAGAATGGGG 240

Qy  262 GATTCACATCTAAGGTTCTGGTCATGGCTGATCAAGGAAGAGAGAGAAATCAGCGAACAAAAAGCC 321
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GATTCACATCTAAGGTTCTGGTCATGGCTGATCAAGGAAGAGAGAGAAATCAGCGAACAAAAAGCC 300

Qy  322 TCTAGGTCTTTCTTACCACAAACACACTCTCTTGCCCACCTGTCTTGAAAAGGGGCAGAAGTA 381
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TCTAGGTCTTTCTTACCACAAACACACTCTCTTGCCCACCTGTCTTGAAAAGGGGCAGAAGTA 360
```

```
Qy  382 TAGTGGGCGAG 392
         |||||||||||
Db  361 TAGTGGGCGAG 371

RESULT 4
CA392561/c

LOCUS       CA392561      586 bp    mRNA    linear    EST 06-NOV-2002
DEFINITION  cs26b04.y1 Human Retinal pigment epithelium/choroid cDNA
            (Un-normalized, unamplified): cs Homo sapiens cDNA clone cs26b04
            5', mRNA sequence.
ACCESSION   CA392561
VERSION     CA392561.1  GI:24725399
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 586)
AUTHORS     Wistow,G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE       Expressed sequence tag analysis of human RPE/choroid for the
            NEIBank Project: Over 6000 non-redundant transcripts, novel genes
            and splice variants
JOURNAL     Mol. Vis. 8 (4), 205-220 (2002)
PUBMED      12107410
MEDLINE     22103460
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Seq primer: M13R1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..586
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cs26b04"
                     /tissue_type="RPE/choroid"
                     /dev_stage="Adult"
                     /lab_host="BMDH10B"
                     /clone_lib="Human Retinal pigment epithelium/choroid cDNA
                     (Un-normalized, unamplified):cs"
                     /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
                     eyes (75-80 years old) yielded approximately 600 mg of
                     dissected RPE/choroid tissue. This in turn yielded 340 ug
                     of total RNA and 7 ug of mRNA. A directionally cloned cDNA
                     library in the pCMVSPORT6 vector was constructed at Life
                     Technologies (Rockville, MD; now part of Invitrogen Corp),
                     essentially following the protocols of the Superscript
                     Plasmid System (Invitrogen Corp.
                     <http://www.invitrogen.com/>). The library code
                     designation was cs. For this library, cDNA inserts were
                     cloned into the NotI/MluI sites of the vector. EST
                     analysis was performed on the unamplified library at the
                     NIH Intramural Sequencing Center (NISC)."
ORIGIN

Query Match     42.8%;  Score 366;  DB 14;  Length 586;
Best Local Similarity   100.0%;  Pred. No. 6.7e-52;
Matches  366;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   27 AGTAGAGGATGGGGTTTCACCGGTGTTAGCCAGGATGGTCTCGGAACTCCTGACCTCGTGATC 86
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  586 AGTAGAGGATGGGGTTTCACCGGTGTTAGCCAGGATGGTCTCGGAACTCCTGACCTCGTGATC 527

Qy   87 CGCCCCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACTGCGCCCAGCCGG 146
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  526 CGCCCCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACTGCGCCCAGCCGG 467
```

RESULT 5

```
Qy   147  TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACCATTTGGGAAC  206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466  TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACCATTTGGGAAC  407

Qy   207  TCCCCCCCACGGCCCATAACTAGTCTCGAGAGTTAAGACCAAGAGCAAGAGAATGGGGATTC  266
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406  TCCCCCCCACGGCCCATAACTAGTCTCGAGAGTTAAGACCAAGAGCAAGAGAATGGGGATTC  407

Qy   267  ACATCTAAGGTCTCGGTGATGGCTGATGAAGGAAGGAAGAAATCAGCGAACAAAAGCCTCTAG  326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346  ACATCTAAGGTCTCGGTGATGGCTGATGAAGGAAGGAAGAAATCAGCGAACAAAAGCCTCTAG  347

Qy   327  GTCTTCTTACCACAAACACCTCTCTGCCCACCTGCTTTGAAAGGGGGCCGAGAAGTATAGTG  386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286  GTCTTCTTACCACAAACACCTCTCTGCCCACCTGCTTTGAAAGGGGGCCGAGAAGTATAGTG  227

Qy   387  GGGGAG  392
          ||||||
Db   226  GGCGAG  221
```

LOCUS   CA392562/c

CA392562

LOCUS   CA392562   586 bp   mRNA   linear   EST 06-NOV-2002

DEFINITION   cs26b05.y1 Human Retinal pigment epithelium/choroid cDNA
          {Un-normalized, unamplified): cs Homo sapiens cDNA clone cs26b05
          5', mRNA sequence.

ACCESSION   CA392562

VERSION   CA392562.1 GI:24725401

KEYWORDS   EST.

SOURCE   Homo sapiens (human)

ORGANISM   Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 586)

AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,
          Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.

TITLE   Expressed sequence tag analysis of human RPE/choroid for the
          NEIBank Project: Over 6000 non-redundant transcripts, novel genes
          and splice variants

JOURNAL   Mol. Vis. 8 (4), 205-220 (2002)

MEDLINE   22103460

PUBMED   12107410

COMMENT   Contact: Wistow G
          Section on Molecular Structure and Function
          National Eye Institute
          6/331, NIH, Bethesda, MD 20892-2740, USA
          Tel: 301 402 3452
          Fax: 301 496 0078
          Email: graeme@helix.nih.gov
          Plate: 26 row: b column: 05
          Seq primer: M13RP1 reverse primer (ABI).

FEATURES   Location/Qualifiers

source   1..586
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="cs26b05"
          /tissue_type="RPE/choroid"
          /dev_stage="Adult"
          /lab_host="EMDH10B"
          /clone_lib="Human Retinal pigment epithelium/choroid cDNA
          (Un-normalized, unamplified): cS"
          /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
          eyes (75-80 years old) yielded approximately 600 mg of
          dissected RPE/choroid tissue. This in turn yielded 340 ug
          of total RNA and 7 ug of mRNA. A directionally cloned cDNA
          library in the pCMVSPORT6 vector was constructed at Life
          Technologies (Rockville, MD; now part of Invitrogen Corp),
          essentially following the protocols of the SuperScript
          plasmid System (Invitrogen Corp.
          <http://www.invitrogen.com/>). The library code

ORIGIN

          designation was cs. For this library, cDNA inserts were
          cloned into the NotI/MluI sites of the vector. EST
          analysis was performed on the unamplified library at the
          NIH Intramural Sequencing Center (NISC)."

Query Match   42.2%; Score 366; DB 14; Length 586;
Best Local Similarity 100.0%; Pred. No. 6.7e-52;
Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    27  AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC   86
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   586  AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC  527

Qy    87  CGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCAGCCGG  146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526  CGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCAGCCGG  467

Qy   147  TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACCATTTGGGAAC  206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466  TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACCATTTGGGAAC  407

Qy   207  TCCCCCCCACGGCCCATAACTAGTCTCGAGAGTTAAGACCAAGAGCAAGAGAATGGGGATTC  266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406  TCCCCCCCACGGCCCATAACTAGTCTCGAGAGTTAAGACCAAGAGCAAGAGAATGGGGATTC  347

Qy   267  ACATCTAAGGTCTCGGTGATGGCTGATGAAGGAAGGAAGAAATCAGCGAACAAAAGCCTCTAG  326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346  ACATCTAAGGTCTCGGTGATGGCTGATGAAGGAAGGAAGAAATCAGCGAACAAAAGCCTCTAG  287

Qy   327  GTCTTCTTACCACAAACACCTCTCTGCCCACCTGCTTTGAAAGGGGGCCGAGAAGTATAGTG  386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286  GTCTTCTTACCACAAACACCTCTCTGCCCACCTGCTTTGAAAGGGGGCCGAGAAGTATAGTG  227

Qy   387  GGGGAG  392
          ||||||
Db   226  GGCGAG  221
```

RESULT 6

LOCUS   BQ710765/c

BQ710765

LOCUS   BQ710765   941 bp   mRNA   linear   EST 16-JUL-2002

DEFINITION   AGENCOURT_8484953 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6301113
          5', mRNA sequence.

ACCESSION   BQ710765

VERSION   BQ710765.1 GI:21849664

KEYWORDS   EST.

SOURCE   Homo sapiens (human)

ORGANISM   Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 941)

AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.

TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL   Unpublished (1999)

COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail1.nih.gov
          Tissue Procurement: Dr. Mark Watson
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov

FEATURES   Location/Qualifiers

source   1..941
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="IMAGE:6301113"
          /lab_host="DH10B (phage-resistant)"

Plate: LLCM2516 row: h column: 10
High quality sequence stop: 583.
Clone distribution: MGC clone distribution information can be
Plate: LLCM2516 row: h column: 10

```
Qy   147  TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACCATTTGGGAAC  206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466  TCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCACCTGACCATTTGGGAAC  407

Qy   207  TCCCCCCCACGGCCCATAACTAGTCTCGAGAGTTAAGACCAAGAGCAAGAGAATGGGGATTC  266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406  TCCCCCCCACGGCCCATAACTAGTCTCGAGAGTTAAGACCAAGAGCAAGAGAATGGGGATTC  407

Qy   267  ACATCTAAGGTCTCGGTGATGGCTGATGAAGGAAGGAAGAAATCAGCGAACAAAAGCCTCTAG  326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346  ACATCTAAGGTCTCGGTGATGGCTGATGAAGGAAGGAAGAAATCAGCGAACAAAAGCCTCTAG  287

Qy   327  GTCTTCTTACCACAAACACCTCTCTGCCCACCTGCTTTGAAAGGGGGCCGAGAAGTATAGTG  386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286  GTCTTCTTACCACAAACACCTCTCTGCCCACCTGCTTTGAAAGGGGGCCGAGAAGTATAGTG  227

Qy   387  GGGGAG  392
          ||||||
Db   226  GGCGAG  221
```

```
RESULT 7

ORIGIN
                /clone_lib="NIH_MGC_113"
                /note="Organ: spleen; Vector: pOT87; Site_1: XhoI; Site_2:
                EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGCACGAG(G). Library constructed by Bing Hong in the
                laboratory of Gerald M. Rubin (University of California,
                Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                Superscript II RT (Life Technologies). Note: this is a
                NIH_MGC Library."

Query Match        31.3%;  Score 271;  DB 13;  Length 941;
Best Local Similarity  99.7%;  Pred. No. 1.8e-36;
Matches 321;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy   509 CAAACAGGGCCCAAATGCATTCATTGAGCCAGGGGGAGGCCAAAGGACTCCGGAGGGAGAGG 568
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   585 CAAACAGGGCCCAAAATGCATTCATTGAGCCAGGGGGAGGCCAAAGGACTCCGGAGGGAGAGG 526

Qy   569 CCCAATAAGGCTGGTGGTGCTATTTCCGATCCATAGAGAGAGAGCGAAGAGTGGGCAGGGGCCCTTTT 628
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   585 CCCAATAAGGCTGGCCCCTATTTCCGATCCATAGAGAGAGAGCGAAGAGTGGGCAGGGCCCTTTT 526

Qy   629 GATTAAGTATCATTCTTTGAATGCAAGCTTCAAAAATCGGGTTATGCCAGGGTGAGGAATGAG 688
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   465 GATTAAGTATCATTCTTTGAATGCAAGCTTCAAAAATCGGGTTATGCCAGGGTGAGGAATGAG 406

Qy   689 CAGGACTAACCATCTGGGTGTCATGGCAAGCCTTCCAGGGCGCCACGGACACAGATCC 748
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   689 CAGGACTAACCATCTGGGTGTCATGGCAAGCCTTCCAGGGCGCCACGGACACAGATCC 748

Qy   749 GCAAGAGGCGCTGCGACCCAGCTCTTCGGTGCCAAGCCACTCGGATTTGGAAACCCGGCGGTCTC 808
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   405 CAGGACTAACCACTCGGGTGTCATGGCAGAAGCCTCTCGGCGACCACTCGGAGACAGATCC 346

Qy   809 AAGGGTCAGCTGTGTAGCCTTGA 830
         |||||||||||||||||||||||
Db   285 AAGGGTCAGCTGTGTAGCCTTGA 264

Qy   345 GCAAGAGGCGCTGCAGCCAGCTCGGTGCCCAAGCCACTCGGATTTGAAACCCGGCGGTCTC 286
```

LOCUS   AI493546   392 bp   mRNA   linear   EST 30-MAR-1999
DEFINITION th36b01.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2120329 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AI493546
VERSION   AI493546.1 GI:4394549
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 392)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies catalog #: 11548-013
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 882   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 391.
FEATURES   Location/Qualifiers
     source   1..392
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2120329"

```
RESULT 8

ORIGIN
                /tissue_type="adenocarcinoma"
                /lab_host="DH10B"
                /clone_lib="NCI_CGAP_Pan1"
                /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.72 kb. Life Technologies catalog #:
                11548-013"

Query Match        29.2%;  Score 253;  DB 9;  Length 392;
Best Local Similarity  100.0%;  Pred. No. 3.9e-33;
Matches 253;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     6 TTTTTTTTTTTTTTTTTTTTTTTTAAGTAGGAGGATGGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140 TTTTTTTTTTTTTTTTTTTTTTAAGTAGGAGGATGGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 199

Qy    66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGGTGGGATTACAGGCA 125
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGGTGGGATTACAGGCA 125

Qy   126 TGAGCCACTGCGCCCCAGCCCGGTCTCTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 185
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   200 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGGTGGGATTACAGGCA 259

Qy   186 ACTCACCTGACATTTGGGAACTCCCCCACGGGCCCATTAACTGATCTTGCAGAGGTAAGACC 245
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   260 TGAGCCACTGCGCCCCAGCCCGGTCTCTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 319

Qy   246 AAGAGCAAGAATG 258
         |||||||||||||
Db   320 ACTCACCTGACATTTGGGAACTCCCCCACGGGCCCATTAACTGATCTTGCAGAGGTAAGACC 379

Qy   380 AAGAGCAAGAATG 392
```

LOCUS   BX504260   615 bp   mRNA   linear   EST 04-SEP-2003
DEFINITION DKFZp686B20133_r1 686 (synonym: h1cc3) Homo sapiens cDNA clone
            DKFZp686B20133 5', mRNA sequence.
ACCESSION   BX504260
VERSION   BX504260.1 GI:32030985
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 615)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE   EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT   Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686B20133) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rpd.de.
FEATURES   Location/Qualifiers
     source   1..615
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686B20133"
                     /dev_stage="adult"

ORIGIN

```
Query Match
Best Local Similarity    14.9%;    Score 129;  DB 13;  Length 615;
Matches 129; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      6 TTTTTTTTTTTTTTTTTTTAAGTAGTAGGAGGAGGAGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 TTTTTTTTTTTTTTTTTTTAAGTAGTAGGAGGAGGAGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 73

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 13

Qy    126 TGAGCCACT 134
            |||||||||
Db     12 TGAGCCACT 4
```

RESULT 9
AA613893

LOCUS   AA613893   507 bp mRNA linear EST 18-FEB-1998
DEFINITION n081h09.s1 NCI_CGAP_AA1 Homo sapiens cDNA clone IMAGE:1113281 3'
     similar to gb:M80651 Hum ORF (HUMAN);contains Alu repetitive
     element.; mRNA sequence.
ACCESSION AA613893
VERSION AA613893.1 GI:2466027
KEYWORDS EST.
SOURCE   Homo sapiens
 ORGANISM Homo sapiens
   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 507)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
   Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
   Email: cgapbs-r@mail.nih.gov
   Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
   Emmert-Buck, M.D., Ph.D.
   cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
   Ph.D.
   cDNA Library Arraying: Greg Lennon, Ph.D.
   DNA Sequencing by: Washington University Genome Sequencing Center
   Clone distribution: NCI-CGAP clone distribution information can be
   found through the I.M.A.G.E. Consortium/LLNL at:
   www-bio.llnl.gov/bbrp/image/image.html
   Insert Length: 1887 Std Error: 0.00
   Seq primer: -40m13 fwd. ET from Amersham
   High quality sequence stop: 469.
FEATURES   Location/Qualifiers
 source   1..507
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:1113281"
     /tissue_type="adrenal adenoma"
     /lab_host="SOLR (kanamycin resistant)"
     /clone_lib="NCI_CGAP_AA1"
     /note="Organ: adrenal gland; Vector: Bluescript SK-;
     Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally.
     Primer: Oligo dT. Two pooled bulk adrenal adenomas. 5'
     adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor
     sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' Average insert
     size: 1.6 kb."
ORIGIN

Query Match
Best Local Similarity   8.8%;   Score 76; DB 9; Length 507;
```

---

```
Best Local Similarity  100.0%;    Pred. No. 0.00039;
Matches 76; Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     68 AACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 AACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 262

Qy    128 AGCCACTGCGCCCAGC 143
            ||||||||||||||||
Db    263 AGCCACTGCGCCCAGC 278
```

RESULT 10
AV719687

LOCUS   AV719687   518 bp mRNA linear EST 16-OCT-2000
DEFINITION AV719687 GLC Homo sapiens cDNA clone GLCCBE07 5', mRNA sequence.
ACCESSION AV719687
VERSION AV719687.1 GI:10816839
KEYWORDS EST.
SOURCE   Homo sapiens
 ORGANISM Homo sapiens
   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 518)
AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z.,
   Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z.,
   Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
   Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE   Homo sapiens cDNA GLC clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
   Chinese National Human Genome Center at Shanghai
   351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
   201203, P. R. China
   Tel: 86-21-50801919(ex.45)
   Fax: 86-21-50801922
   Email: hanzg@chgc.sh.cn
   This clone is available at CHGC in Shanghai.
FEATURES   Location/Qualifiers
 source   1..518
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="GLCCBE07"
     /tissue_type="corresponding non cancerous liver tissue"
     /dev_stage="Adult"
     /lab_host="SOLR"
     /clone_lib="GLC"
     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
     XhoI"
ORIGIN

Query Match
Best Local Similarity   8.8%;   Score 76; DB 9; Length 518;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   68 AACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   216 AACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 275

Qy   128 AGCCACTGCGCCCAGC 143
            ||||||||||||||||
Db   276 AGCCACTGCGCCCAGC 291
```

RESULT 11
AV718792

LOCUS   AV718792   592 bp mRNA linear EST 16-OCT-2000
DEFINITION AV718792 GLC Homo sapiens cDNA clone GLCCOE12 5', mRNA sequence.
ACCESSION AV718792
VERSION AV718792.1 GI:10815944
KEYWORDS EST.
SOURCE   Homo sapiens (human)
```

ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 592)
AUTHORS    Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z.,
           Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z.,
           Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
           Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
           Homo sapiens cDNA GLC clones
TITLE      Unpublished (2000)
JOURNAL
COMMENT    Contact: Zeguang Han
           Chinese National Human Genome Center at Shanghai
           351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
           201203, P. R. China
           Tel: 86-21-50801919(ex.45)
           Fax: 86-21-50801922
           Email: hanzg@chgc.sh.cn
           This clone is available at CHGC in Shanghai.
FEATURES
     source          1..592
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="GLCCOB12"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /clone_lib="GLC"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

Query Match     8.8%;   Score 76;  DB 9;  Length 592;
Best Local Similarity   100.0%;   Pred. No. 0.00034;
Matches  76;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   68 AACTCCTGACCTCGTGATCCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  191 AACTCCTGACCTCGTGATCCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 250

Qy  128 AGCCACTGCGCCCAGC 143
        ||||||||||||||||
Db  251 AGCCACTGCGCCCAGC 266

RESULT 12
LOCUS       AV719326          601 bp    mRNA    linear    EST 16-OCT-2000
DEFINITION  AV719326 GLC Homo sapiens cDNA clone GLCCNF05 5', mRNA sequence.
ACCESSION   AV719326
VERSION     AV719326.1   GI:10816478
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 601)
  AUTHORS   Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z.,
            Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z.,
            Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
            Homo sapiens cDNA GLC clones
  TITLE     Unpublished (2000)
  JOURNAL
  COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES
     source          1..601
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="GLCCNF05"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /clone_lib="GLC"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

Query Match     8.8%;   Score 76;  DB 9;  Length 601;
Best Local Similarity   100.0%;   Pred. No. 0.00033;
Matches  76;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   68 AACTCCTGACCTCGTGATCCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 AACTCCTGACCTCGTGATCCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 258

Qy  128 AGCCACTGCGCCCAGC 143
        ||||||||||||||||
Db  259 AGCCACTGCGCCCAGC 274

RESULT 13
LOCUS       AV720014          604 bp    mRNA    linear    EST 16-OCT-2000
DEFINITION  AV720014 GLC Homo sapiens cDNA clone GLCCNG04 5', mRNA sequence.
ACCESSION   AV720014
VERSION     AV720014.1   GI:10817166
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 604)
  AUTHORS   Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z.,
            Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z.,
            Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
            Homo sapiens cDNA GLC clones
  TITLE     Unpublished (2000)
  JOURNAL
  COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES
     source          1..604
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="GLCCNG04"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /clone_lib="GLC"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

Query Match     8.8%;   Score 76;  DB 9;  Length 604;
Best Local Similarity   100.0%;   Pred. No. 0.00033;
Matches  76;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   68 AACTCCTGACCTCGTGATCCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  206 AACTCCTGACCTCGTGATCCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 265

```
Qy      128 AGCCACTGCGCCCAGC 143
            |||||||||||||||||
Db      266 AGCCACTGCGCCCAGC 281
```

RESULT 14
LOCUS   AV720842
DEFINITION   AV720842 GLC Homo sapiens cDNA clone GLCCOH05 5', mRNA sequence.
ACCESSION   AV720842
VERSION   AV720842.1 GI:10817994
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 613)
  AUTHORS   Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z.,
              Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z.,
              Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
              Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
  TITLE   Homo sapiens cDNA GLC clones
  JOURNAL   Unpublished (2000)
  COMMENT   Contact: Zeguang Han
              Chinese National Human Genome Center at Shanghai
              351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
              201203, P. R. China
              Tel: 86-21-50801919(ex.45)
              Fax: 86-21-50801922
              Email: hanzg@chgc.sh.cn
              This clone is available at CHGC in Shanghai.
FEATURES   Location/Qualifiers
     source   1..613
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="GLCCOH05"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /clone_lib="GLC"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

Query Match   8.8%; Score 76; DB 9; Length 613;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       68 AACTCCTGACCTCGTGATCCGGCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      187 AACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 246

Qy      128 AGCCACTGCGCCCAGC 143
            ||||||||||||||||
Db      247 AGCCACTGCGCCCAGC 262
```

RESULT 15
LOCUS   BM561574
DEFINITION   BM561574 AGENCOURT_6567251 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5739759
              5', mRNA sequence.
ACCESSION   BM561574
VERSION   BM561574.1 GI:18806941
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 481)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)

[second column]

JOURNAL
COMMENT   Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM12753 row: j column: 16
              High quality sequence stop: 480.
FEATURES   Location/Qualifiers
     source   1..481
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5739759"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_88"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
ORIGIN

Query Match   8.5%; Score 74; DB 12; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       60 TGGTCTCGAACTCCTGACCTCGTGATCCGGCCGCCTCGGCCTCCCAAAGTGCTGGGATTA 119
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22 TGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTA 81

Qy      120 CAGGCATGAGCCAC 133
            ||||||||||||||
Db       82 CAGGCATGAGCCAC 95
```

Search completed: May 24, 2004, 18:18:40
Job time : 2344 secs

```
                                                    GenCore version 5.1.6
                                          Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          May 24, 2004, 13:20:05 ; Search time 333 Seconds
                                                       (without alignments)
                                                     11060.624 Million cell updates/sec

Title:           US-10-621-363-12
Perfect score:   867
Sequence:        1 tcgggtttttttttttttttt.........accaatctcgtcgcgaattc 867

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched:        3373863 seqs, 2124059041 residues

Total number of hits satisfying chosen parameters:        6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_29Jan04:*
                 1: geneseqm1980s:*
                 2: geneseqm1990s:*
                 3: geneseqn2000s:*
                 4: geneseqn2001as:*
                 5: geneseqn2001bs:*
                 6: geneseqn2002s:*
                 7: geneseqn2003as:*
                 8: geneseqn2003bs:*
                 9: geneseqn2003cs:*
                10: geneseqn2004s:*

                                          SUMMARIES

Result          Query                     %
No.     Score  Match Length  DB ID       Description
------  -----  ------------  -----        -----------
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
Result          Query                     %
No.     Score  Match Length  DB ID       Description
 1      865.2  99.8    867   2  AAX30367  Aax30367 DNA encod
 2      865.2  99.8    867   7  ADB47729  Adb47729 Novel hum
 3      822.6  94.9  58181   7  ABZ68140  Abz68140 Human sec
 4      822.6  94.9  58181   7  ABZ74619  Abz74619 Secreted
 5      822.6  94.9  58181   9  ADC21010  Adc21010 Human sec
 6      133.6  15.4  39703   9  ADC86430  Adc86430 Human GPC
 7      133.6  15.4 134292   9  ACA64895  Aca64895 Human GAB
 8      130.4  15.0  22255   5  AAH26497  Aah26497 Human low
 9      130.4  15.0 115756   7  ACD13448  Acd13448 Human DNA
10      129.4  14.9  20951   7  AAL53468  Aal53468 Genomic D
11      128.8  14.9  94191   9  ADE11169_3 Continuation (4 of
12      128    14.8    407   5  ABV12806  Abv12806 Human pro
13      128    14.8    447   5  ABV03637  Abv03637 Human pro
14      128    14.8    458   5  ABV33935  Abv33935 Human pro
15      128    14.8    458   5  ABV42808  Abv42808 Human pro
16      127.8  14.7    307   4  AAK84065  Aak84065 Human imm
17      127.8  14.7   5069   4  AAK67404  Aak67404 Human imm
18      127.8  14.7  51959   6  ABQ99650  Abq99650 Human mem
19      127.8  14.7   8960   6  AAK69657  Aak69657 Human imm
20      127.2  14.7    172   4  AAK81214  Aak81214 Human imm
21      127    14.6    484   5  ABV57673  Abv57673 Human pro
22      126.6  14.6 121600   6  ABT10748  Abt10748 Human bre
23      126.6  14.6    486   5  ABV49991  Abv49991 Human pro
24      126.6  14.6    506   4  AAK86883  Aak86883 Human imm
25      126.2  14.6    304   4  AAK84486  Aak84486 Human imm
26      126.2  14.6    304   4  AAL05753  Aal05753 Human rep
27      126.2  14.6    417   4  AAH10519  Aah10519 Human cDN
28      126.2  14.6   2359   4  AAH18461  Aah18461 Human cDN
29      126.2  14.6 149671   6  ABK84797  Abk84797 Human cDN
30      126.2  14.6 149671   8  ADB70361  Adb70361 Moesin cD
31      126    14.5    932   4  AAK84143  Aak84143 Human imm
32      126    14.5    932   4  AAK84144  Aak84144 Human imm
33      126    14.5   7152   5  AAD26830  Aad26830 Human SNA
34      126    14.5  36797   6  AAD26738  Aad26738 Human syn
35      125.8  14.5    380   6  ABL86703  Abl86703 Human ova
36      125.8  14.5 136284   6  ABK83575  Abk83575 Human cDN
37      125.6  14.5   5409   4  AAS28318  Aas28318 Genomic s
38      125.6  14.5  11556   6  ABK69902  Abk69902 Human sec
39      125.4  14.5  31952   4  AAK89370  Aak89370 Human dig
40      125.4  14.5   5540   4  AAK79344  Aak79344 Human imm
41      125.4  14.5  13836   5  ABA16633  Aba16633 Human ner
42      125.4  14.5  17197   4  AAK79870  Aak79870 Human imm
43      125.4  14.5  30420   4  AAL36067  Aal36067 Human mus
44      125.4  14.5  30420   4  AAL06476  Aal06476 Human rep
45      125.4  14.5  30420   5  AAS40588  Aas40588 DNA encod
```

                                          ALIGNMENTS

```
RESULT 1
AAX30367

ID   AAX30367 standard; DNA; 867 BP.
XX
AC   AAX30367;
XX
DT   14-MAY-1999  (first entry)
XX
DE   DNA encoding a human secreted protein.
XX
KW   Secreted protein; cancer; tumour; neurodegenerative disorder;
KW   developmental abnormality; foetal deficiency; blood disorder;
KW   CNS disorder; immune system disease; autoimmune disease; hepatic disease;
KW   renal disease; diabetes; inflammation; allergy; ischemic shock;
KW   Alzheimer's; cognitive disorder; schizophrenia; cardiovascular disorder;
KW   prostate disease; asthma; osteoporosis; arthritis; ss.
XX
OS   Homo sapiens.
XX
PN   WO9907891-A1.
XX
PD   18-FEB-1999.
XX
PF   04-AUG-1998;              98WO-US016235.
XX
PR   05-AUG-1997;              97US-0054798P.
PR   05-AUG-1997;              97US-0054803P.
PR   05-AUG-1997;              97US-0054804P.
PR   05-AUG-1997;              97US-0054806P.
PR   05-AUG-1997;              97US-0054807P.
PR   05-AUG-1997;              97US-0054808P.
PR   05-AUG-1997;              97US-0054809P.
PR   05-AUG-1997;              97US-0053309P.
PR   05-AUG-1997;              97US-0053310P.
PR   05-AUG-1997;              97US-0053311P.
PR   05-AUG-1997;              97US-0053312P.
PR   05-AUG-1997;              97US-0053386P.
PR   18-AUG-1997;              97US-0055970P.
PR   18-AUG-1997;              97US-0055986P.
PR   19-AUG-1997;              97US-0056006P.
PR   19-AUG-1997;              97US-0056365P.
PR   19-AUG-1997;              97US-0056366P.
PR   19-AUG-1997;              97US-0056367P.
PR   19-AUG-1997;              97US-0056370P.
PR   19-AUG-1997;              97US-0056371P.
PR   19-AUG-1997;              97US-0056557P.
```

PR 19-AUG-1997; 97US-0056563P.
PR 19-AUG-1997; 97US-0056731P.
PR 19-AUG-1997; 97US-0056732P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Soppet DR, Ebner R, Olsen HS, Young PE, Greene JM;
PI Ferrie AM, Yu G, Ni J, Rosen CA, Brewer LA, Janat F;
XX
DR WPI; 1999-167452/14.
DR P-PSDB; AAY10846.
XX
PT New isolated human genes encoding secreted polypeptides - useful for
PT diagnosis and treatment of pathological diseases.
XX
PS Claim 3; Page 253-254; 331pp; English.
XX
CC The specification describes secreted proteins and their corresponding
CC polynucleotides which are useful for preventing, treating or ameliorating
CC medical conditions, e.g. by protein or gene therapy. Pathological
CC medical conditions can also be diagnosed by determining the amount of the
CC secreted polypeptides in a sample or by determining the presence of
CC conditions in the polynucleotides. Specific uses are described for each of
CC mutations based on which tissues they are most highly expressed in,
CC the products, based on which tissues they are most highly expressed in,
CC and include developing products for the diagnosis or treatment of cancer,
CC tumours, neurodegenerative disorders, developmental abnormalities and
CC foetal deficiencies, blood disorders, CNS disorders, diseases of the
CC immune system, autoimmune diseases, hepatic and renal disease, diabetes,
CC inflammation, allergies, ischemic shock, Alzheimer's and cognitive
CC disorders, schizophrenia, cardiovascular disorders, prostate diseases,
CC asthma, disorders involving osteoclasts such as osteoporosis, arthritis
CC or malignancies, diseases of testes, lung or thymus, digestive/endocrine
CC disorders, infections and AIDS. The polypeptides are also useful for
CC identifying their binding partners
XX
SQ Sequence 867 BP; 219 A; 231 C; 222 G; 192 T; 0 U; 3 Other;

Query Match 99.8%; Score 865.2; DB 2; Length 867;
Best Local Similarity 100.0%; Pred. No. 2.1e-247;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTTCACCGTGTTAGCCAGGAT 60
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTTCACCGTGTTAGCCAGGAT 60

Qy 61 GGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTAC 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTAC 120

Qy 121 AGGCATGAGCCACTCGCGCCCAGCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAA 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGGCATGAGCCACTCGCGCCCAGCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAA 180

Qy 181 CCCATACTCACCTGACATTTGGGAACTCCCCCCACGGCCATAACTGAICTGCAGAGGTA 240
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CCCATACTCACCTGACATTTGGGAACTCCCCCCACGGCCATAACTGAICTGCAGAGGTA 240

Qy 241 AGACCAAGAGCCAAGAGATGGGGGATTCACATCTAAGGTCTGGTGATGAGGAAG 300
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 AGACCAAGAGCCAAGAGATGGGGGATTCACATCTAAGGTCTGGTGATGAGGAAG 300

Qy 301 AAGAATCAGCGAACAAAAAGCCTCTAGGTCTTTTCTTACCACAAAACACCTCTCTGCCCACCT 360
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 AAGAATCAGCGAACAAAAAGCCTCTAGGTCTTTTCTTACCACAAAACACCTCTCTGCCCACCT 360

Qy 361 GCTTTGAAAGGGGCAGAAGTATAGTGGGCGAGCTCCCCACCTGCTTACAGTGAAGGGATCT 420
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 GCTTTGAAAGGGGCAGAAGTATAGTGGGCGAGCTCCCCACCTGCTTACAGTGAAGGGATCT 420

Qy 421 GGAGAAATACTCACACTTTGAGGTGCCGCCCTCTTCATCAGCCAGCTCTAACTTAAGCC 480
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GGAGAAATACTCACACTTTGAGGTGCCGCCCTCTTCATCAGCCAGCTCTAACTTAAGCC 480

Qy 481 AATGAACCCCACCGGGAGCTTACACAAGTYCAAAAACGAGTCAAATGCAATCCATCAGGG 540
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 AATGAACCCCACCGGGAGCTTACACAAGTYCAAAAACGAGTCAAATGCAATCCATCAGGG 540

Qy 541 GAGGCCAAAGGACTCCCGGAGGAGAGAGGCCCAATAAAGGCTGGTGCTATTTCCGATCCATA 600
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GAGGCCAAAGGACTCCCGGAGGAGAGAGGCCCAATAAAGGCTGGTGCTATTTCCGATCCATA 600

Qy 601 GAGAGAGCAGGAGTGGGGCCAGGCCCTTTGATTAATGTATCATTCTTGAATGCAAGCTTCA 660
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 GAGAGAGCAGGAGTGGGGCCAGGCCCTTTGATTAATGTATCATTCTTGAATGCAAGCTTCA 660

Qy 661 AAATCCGGGTATGCCGGGTGAGAATGAGCAGGACTAACACCTGGGTGTGTCATGGCAAGCCT 720
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 AAATCCGGGTATGCCGGGTGAGAATGAGCAGGACTAACACCTGGGTGTGTCATGGCAAGCCT 720

Qy 721 CCAGGGCCCGACTGGCCAGAGACAGATCCGCAAGAGGCTCTGCAGCCAGTCTCTGGTGCCAA 780
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 CCAGGGCCCGACTGGCCAGAGACAGATCCGCAAGAGGCTTCGCAGCCAGTCTCGGTGCCAA 780

Qy 781 GCCACTCCGGATTTGAACCCCCGGCTCCTCAAGGTGTCAGGCTGTGTTAGCCCTTGANTGAAYCACC 840
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 GCCACTCCGGATTTGAACCCCCGGCTCCTCAAGGTGTCAGGCTGTGTTAGCCCTTGANTGAAYCACC 840

Qy 841 TGCTATGACCAATTCGTNCCCGAATTC 867
   |||||||||||||||||||||||||||
Db 841 TGCTATGACCAATTCGTNCCCGAATTC 867

RESULT 2
AD847729
ID AD847729 standard; cDNA; 867 BP.
XX
AC AD847729;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human secreted protein cDNA #2.
XX
KW ss; gene; human; secreted protein; insulin; haemoglobin S; haemoglobin B;
KW superoxidase; SOD; catalase; DNA repair protein; oncogene;
KW tumour suppressor; tumour necrosis factor; TNF; inflammation;
KW blood vessel growth inhibition; immune response; immune system disorder;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW peripheral artery disease; limb ischaemia; arterio-arterial fistula;
KW arteriovenous fistula; congenital heart defect;
KW neovascularisation disorder; wound healing;
KW epithelial cell proliferation; neurological disease; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; mania; dementia;
KW infectious disease.
XX
OS Homo sapiens.
XX
PN US2003054443-A1.
XX
PD 20-MAR-2003.
XX
PF 04-OCT-2001; 2001US-00969730.
XX
PR 05-AUG-1997; 97US-0054798P.
PR 05-AUG-1997; 97US-0054803P.
PR 05-AUG-1997; 97US-0054804P.
PR 05-AUG-1997; 97US-0054806P.
PR 05-AUG-1997; 97US-0054807P.
PR 05-AUG-1997; 97US-0054808P.
PR 05-AUG-1997; 97US-0054809P.
PR 05-AUG-1997; 97US-0053109P.
PR 05-AUG-1997; 97US-0053110P.
PR 05-AUG-1997; 97US-0053311P.
PR 05-AUG-1997; 97US-0053312P.
PR 18-AUG-1997; 97US-0053386P.
PR 18-AUG-1997; 97US-0055970P.
PR 18-AUG-1997; 97US-0055986P.

```
PR   19-AUG-1997;  9TUS-0056364P.
PR   19-AUG-1997;  9TUS-0056365P.
PR   19-AUG-1997;  9TUS-0056366P.
PR   19-AUG-1997;  9TUS-0056367P.
PR   19-AUG-1997;  9TUS-0055370P.
PR   19-AUG-1997;  9TUS-0053371P.
PR   19-AUG-1997;  9TUS-0055557P.
PR   19-AUG-1997;  9TUS-0055563P.
PR   19-AUG-1997;  9TUS-0055731P.
PR   19-AUG-1997;  9TUS-0056732P.
PR   04-AUG-1998;  98WO-US016235.
PR   04-FEB-1999;  99US-0024112.
PR   06-OCT-2000;  2000US-0282912.
PR   01-FEB-2001;  2001US-00774639.
XX
PA   (RUBE/)  RUBEN S M.
PA   (SOPP/)  SOPPET D R.
PA   (EBNE/)  EBNER R.
PA   (OLSE/)  OLSEN H S.
PA   (YOIN/)  YOUNG P E.
PA   (GREE/)  GREENE J M.
PA   (FERR/)  FERRIE A M.
PA   (YUGG/)  YU G.
PA   (NIGJ/)  NI J.
PA   (ROSE/)  ROSEN C A.
PA   (BREW/)  BREWER L A.
PA   (JANA/)  JANAT F.
PA   (BIRS/)  BIRSE C E.
XX
PI   Ruben SM, Soppet DR, Ebner R, Olsen HS, Young PE, Greene JM;
PI   Ferrie AM, Yu G, Ni J, Rosen CA, Brewer LA, Janat F, Birse CE;
XX
DR   W21; 2003-699903/66.
DR   P-PSDB; ADB47826.
XX
PT   Novel human secreted proteins useful for treating and/or diagnosing
PT   disorders of immune system, cardiovascular disorders such as peripheral
PT   artery disease, neurological diseases such as Alzheimer's disease.
XX
PS   Claim 3; Page 205; 333pp; English.
XX
CC   The invention relates to novel human secreted proteins. The protein is
CC   useful for preventing, treating or ameliorating a medical condition. The
CC   protein is useful for diagnosing a pathological condition or
CC   susceptibility to a pathological condition in a subject. The protein is
CC   useful for identifying a binding partner. The nucleic acid is useful for
CC   diagnosing pathological condition or a susceptibility to pathological
CC   condition in a subject. The protein is useful as reagents for
CC   differential identification of the tissues or cell types present in a
CC   biological sample. The protein can be administered to patients having
CC   absent or decreased levels of polypeptides e.g. insulin, to supplement
CC   absent or decreased levels of different polypeptides, e.g. haemoglobin s
CC   for haemoglobin B, superoxidase (SOD), catalase, DNA repair protein, to
CC   inhibit the activity of a polypeptide e.g. an oncogene or tumour
CC   suppressor, to activate the activity of polypeptide e.g. by binding to a
CC   receptor, to reduce the activity of membrane bound receptor by competing
CC   with it for free ligand e.g. soluble tumour necrosis factor (TNF)
CC   receptors used in reducing inflammation, or to bring about a desired
CC   response e.g. blood vessel growth inhibition, enhancement of immune
CC   response to proliferative cells or tissues. The protein and the nucleic
CC   acid are useful for treating, preventing, detecting, diagnosing disorders
CC   of immune system involving abnormal growth of specific types of cells as
CC   well as of other cell types where expression has been observed. The
CC   protein, the nucleic acid and antibodies are useful for treating,
CC   preventing and/or diagnosing diseases, disorders and/or conditions of
CC   immune system, hyperproliferative disorders including neoplasms,
CC   cardiovascular disorders (such as peripheral artery disease, limb
CC   ischaemia, arterio-arterial fistula, arteriovenous fistula, congenital
CC   heart defects, etc), neovascularisation disorders, wound healing and
CC   epithelial cell proliferation, neurological diseases (such as Alzheimer's
CC   disease, Parkinson's disease, Huntington's disease, mania, dementia,
CC   etc), infectious diseases caused by virus, bacteria, fungi, etc. The
CC   present sequence represents a cDNA encoding a novel human secreted
```

```
CC   protein.
XX
SQ   Sequence 867 BP; 219 A; 231 C; 222 G; 192 T; 0 U; 3 Other;

     Query Match             99.8%;  Score 865.2;  DB 9;  Length 867;
     Best Local Similarity   100.0%;  Pred. No. 2.1e-247;
     Matches  867;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGGTTTTCACCGTGTTAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGGTTTTCACCGTGTTAG 60

Qy         61 GGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTAC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTAC 120

Qy        121 AGGCATGAGCCACTCCGCCCAGCCGGTCTCTTTTAAACATTCCCCAGGGACTGTACAGCCAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCATGAGCCACTCCGCCCAGCCGGTCTCTTTTAAACATTCCCCAGGGACTGTACAGCCAA 180

Qy        181 CCCATACTCACCTGACATTTGGGAACTCCCCCCACGGGCCCATAACTGATCTGCAGAGGTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCATACTCACCTGACATTTGGGAACTCCCCCCACGGGCCCATAACTGATCTGCAGAGGTA 240

Qy        241 AGACCAAGAGCCAAGAAATGGGGGGATTCACATCTAAGGTCTGGTGATGGGCTGATGAGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGACCAAGAGCCAAGAAATGGGGGGATTCACATCTAAGGTCTGGTGATGGGCTGATGAGGAG 300

Qy        301 AAGAATCAGCGAACAAAAAGCCTCTAGGTCTTTCTTCTTACCACAAACACCTCTCTGCCCACCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGAATCAGCGAACAAAAAGCCTCTAGGTCTTTCTTCTTACCACAAACACCTCTCTGCCCACCT 360

Qy        361 GCTTTGAAAGGGGCAGAAGTATAGTGGGCGGAGCTGCCCACCTGTACAGTGAAGGGATCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTTGAAAGGGGCAGAAGTATAGTGGGCGGAGCTGCCCACCTGTACAGTGAAGGGATCT 420

Qy        421 GGAGAAATACTCACACTTTGAGGTGTCTCGCCCTCTTCATCAGCCAGCCTCTAACTTAAGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAGAAATACTCACACTTTGAGGTGTCTCGCCCTCTTCATCAGCCAGCCTCTAACTTAAGCC 480

Qy        481 AATGACCCCACGGAGAGCTTACACAAGTYCAAACAGGGCCCAAATGCATTCATGAGCAGGGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATGACCCCACGGAGAGCTTACACAAGTYCAAACAGGGCCCAAATGCATTCATGAGCAGGGG 540

Qy        541 GAGGGCCAAACGGACTCCGGGAGGAGGAGGCCAATAAGGCCCCATAAACTGATCATCCATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGCCAAACGGACTCCGGGAGGAGGAGGCCAATAAGGCCCCATAAACTGATCATCCATA 600

Qy        601 GAGAGAGGCAGAGGTGGGCAGGCCCCTTTGAATTAAATGTATCATTCTTGAAATGCAAGCTTCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAGAGGCAGAGGTGGGCAGGCCCCTTTGAATTAAATGTATCATTCTTGAAATGCAAGCTTCA 660

Qy        661 AAATCCGGGTTATGCCGGGGTGAGATGAGCAGGACTAAACACCTGGGTGTCATGGCAAGCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAATCCGGGTTATGCCGGGGTGAGATGAGCAGGACTAAACACCTGGGTGTCATGGCAAGCCT 720

Qy        721 CCAGGGCCGACTGGCCCAGAGACAGATCCGGCAAGAGGCCTCTGGTCAGCCAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCAGGGCCGACTGGCCCAGAGACAGATCCGGCAAGAGGCCTCTGGTCAGCCAA 780

Qy        781 GCCACCTCGGATTTGAACCCCGGGCTCCCTGGGTGTTGTAGCCTTGANTGAAYCACC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCACCTCGGATTTGAACCCCGGGCTCCCTGGGTGTTGTAGCCTTGANTGAAYCACC 840

Qy        841 TGCTATGACCAATCTCGTGCCGGAATTC 867
              |||||||||||||||||||||||||||||
Db        841 TGCTATGACCAATCTCGTGCCGGAATTC 867

XX
RESULT 3
AB268140/c
ID   AB268140 standard; DNA; 58181 BP.
XX
```

AC AB268140;

XX

DT 26-MAR-2003 (first entry)

XX

DE Human secreted protein encoding genomic DNA SEQ ID NO 1663.

XX

KW Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnerary; antibacterial; antiparkinsonian; antisickling; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antiulcer; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.

XX

OS Homo sapiens.

XX

PN WO20027186-A2.

XX

PD 03-OCT-2002.

XX

PF 26-MAR-2002; 2002WO-US009188.

XX

PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ruben SM;

DR WPI; 2003-040583/03.

XX

FS Disclosure; Page 2272-2286; 2423pp; English.

XX

PT The invention relates to novel human genes (AB266891-AB268209) and the
PT encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
PT treating or ameliorating medical conditions e.g. by protein or gene
PT therapy. The genes are isolated from a range of human tissues disclosed
PT in the specification. The nucleic acids, proteins, antibodies and
PT (ant)agonists are useful in the diagnosis, treatment and prevention of:
PT (a) cancer, e.g. breast and ovarian cancer and other cancers of the
PT adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
PT lung or urogenital; (b) immune disorders e.g. Addison's disease,
PT allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
PT diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
PT arthritis and ulcerative colitis; (c) cardiovascular disorders such as
PT myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
PT cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
PT bacterial, fungal and parasitic infections

XX

SQ Sequence 58181 BP; 15503 A; 12755 C; 13504 G; 16419 T; 0 U; 0 Other;

Query Match 94.9%; Score 822.6; DB 7; Length 58181;
Best Local Similarity 99.3%; Pred. No. 6.2e-234;
Matches 844; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

Qy 6 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
  |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db 52565 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 52506

Qy 66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52505 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 52446

Qy 126 TGAGCCACTGCGCCCAGCCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 185
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52445 TGAGCCACTGCGCCCAGCCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 52386

Qy 186 ACTTCACCTGACATTTGGGAACTCCCCCCCCCCACGGCCCATRACTGATCTGCAGAGGTAAGACC 245
  ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52385 ACTTCACCTGACATTTGGGAACTCCCCCCCCACGGCCCATRACTGATCTGCAGAGGTAAGACC 52326

Qy 246 AAGAGCAAGAAATGGGGGAGGATTCACATTCAAAGGTCTGGTGGATGGCATGGATGCTGAGAAGGAAGAAGAA 305
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52325 AAGAGCAAGAAATGGGGGAGGATTCACATTCAAAGGTCTGGTGGATGGCATGGATGCTGAGAAGGAAGAAGAA 52266

Qy 306 TCAGCGGACAAAAGCCTCTAGGTGCTTTTCTTACCACAAACACACTTCTCTCGCCCACCTGCTTT 365
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52265 TCAGCGGACAAAAGCCTCTAGGTCTTTTCTTACCACAAACACACTTCTCTCGCCCACCTGCTTT 52206

Qy 366 GAAAGGGCGCAGAAGTATTAGTGGGCGGA-GCTGCCCACCTTGCTACAGTGAAAGGGGATTCTGGAG 424
  |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 52205 GAAAGGGCGCAGAAGTATTAGTGGGCGGAGCTGCCCACCTTGCTACAGTGAAAGGGGATTCTGGAG 52146

Qy 425 AAATACTCACACTTTGAGGTGCTTCGGCCCCTTCTCAATCAGGCCAGTCTCTTAAGCCAATG 484
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52145 AAATACTCACACTTTGAGGTGCTTCGGCCCCTTCTCAATCAGGCCAGTCTCTTAAGCCAATG 52086

Qy 485 ACCCCACGGGAG-CTTACACAAGTYCAAACAGGCCCAAATGCAATTCATGAGCAGGGGAG 543
  |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db 52085 ACCCCACGGGAGCCTTACACAAGTYCAAACAGGCCCAAATGCAATTCATGAGCAGGGGAG 52026

Qy 544 GCCAAAAGACTCCGGAGAGAGAGGGCCCAATAAGGTCGTTGCTATTTCCGATCCATAGAG 603
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 52025 GCCAAAAGACTCCGGAGAGAGAGGGCCCAATAAGGTCGTTGCTATTTCCGATCCATAGAG 51966

Qy 604 AGAGCAGAGGTGGCAGGCGCCCTTTGATTAAATGTATCATTCTTGAAATGCAAGCTTCAAA 663
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 51965 AGAGCAGAGGTGGCAGGCGCCCTTTGATTAAATGTATCATTCTTGAAATGCAAGCTTCAAA 51906

Qy 664 TCCGGGTATGCCGGGGTGAGAATGAGGACAGGACTAACACTGGGGTTGTCATGGCAAGCCTGC 723
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 51905 TCCGGGTATGCCGGGGTGAGAATGAGGACAGGACTAACACTGGGGTTGTCATGGCAAGCCTCA 51846

Qy 724 GGGGCGACTGGCCAGAGACAGAGATCGCGAAGAGAGGCTTCTGCAGCCAGCCTCTGGGGTGCCAAGCC 783
  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 51785 ACTCGGAGTTGAACCCGGGCCTTCTTCAAGGTCAGTGTGTGTAGCCTTGAAANGAAATCACCTGC 51726

Qy 844 TATGACCAAT 853
  ||||||||||
Db 51725 TATGACCAAT 51716

RESULT 4

ID AB274619 standard; DNA; 58181 BP.

XX

AC AB274619;

XX

DT 12-MAY-2003 (first entry)

XX

DE Secreted protein gene 372 genomic fragment HWBCN336, SEQ ID NO:1766.

XX

KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW drug screening; chromosome identification; chromosome mapping;
KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW antianaemic; vulnerary; gene; ds.

XX

OS Homo sapiens.

XX

PN WO20027013-A2.

XX

PD 03-OCT-2002.

XX

```
PF   26-MAR-2002; 2002WO-US009370.
XX
PR   27-MAR-2001; 2001US-0278650P.
PR   12-SEP-2001; 2001US-00950082.
PR   12-SEP-2001; 2001US-00950083.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,    Ruben SM;
XX
DR   WPI; 2003-040578/03.
XX
PT   New human secreted proteins and nucleic acids, useful for detecting or
PT   treating cancer or other hyperproliferative disorders, autoimmune
PT   disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
XX
PS   Disclosure; Page 2331-2345; 2474pp; English.
XX
CC   AB273281-AB273697 represent cDNAs corresponding to 391 human secreted
CC   protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC   AB273698-AB274687 represent human secreted protein genomic fragments. The
CC   invention also encompasses antibodies specific for the secreted proteins,
CC   the use of the secreted proteins in drug screening and recombinant
CC   vectors and host cells comprising a nucleic acid of the invention. The
CC   secreted proteins are thought to be involved in biological activities
CC   associated with cellular signaling, cellular differentiation, cell
CC   migration, prohormone activation and neurotransmitter activity. The
CC   secreted proteins, nucleic acids encoding them, antibodies or antibody
CC   fragments specific for the secreted proteins, and modulators of protein
CC   activity are useful for diagnosing or treating cancers or other
CC   hyperproliferative disorders. Additionally, the secreted proteins and
CC   their nucleic acids may also be used in the treatment of autoimmune
CC   disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC   (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC   wound healing. Nucleic acids of the invention may be used for chromosome
CC   identification, chromosome mapping, in gene therapy, for identifying
CC   individuals from minute biological samples, as hybridisation probes, and
CC   as molecular weight markers. The present sequence represents a human
CC   secreted protein genomic fragment referred to in the disclosure of the
CC   invention
XX
SQ   Sequence 58181 BP; 15503 A; 12755 C; 13504 G; 16419 T; 0 U; 0 Other;

Query Match             94.9%;  Score 822.6;  DB 7;  Length 58181;
Best Local Similarity   99.3%;
Matches  844;  Conservative  2;  Pred. No. 6.2e-234;
                               Mismatches  2;  Indels  2;  Gaps  2;

Qy      6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52565 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 52506

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
          |||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db  52505 CGAACTCCTGACCTCCTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCA 52446

Qy    126 TGAGCCACTGCGCCCAGCCGGTCTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 185
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  52445 TGAGCCACTGCGCCCAGCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 52386

Qy    186 ACTCACCTGACATTTGGGAACTCCCCCCCCAGGGCCATAACTGATCTGCAGAGGTAAGACC 245
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  52385 ACTCACCTGACATTTGGGAACTCCCCCACGGCCATAACTGATCTGCAGAGGTAAGACC 52326

Qy    246 AAGAGCAAAGAAATGGGGGGATTCACATCTAAGGTCTTGGTTGATGGGCTCATGAGGGAGAGAA 305
          |||||||||||||||||||||||||||||||||||||| ||| ||||| |||||||  ||||||
Db  52325 AAGAGCAAAGAAATGGGGGGATTCACATCTAAGGTCTGGGTGATGGCTCATGAGGAAGAGAA 52266

Qy    306 TCAGCGGAACAAAAAGCCTTCTAGGTCTTTCTTACCACAAAACAACTTCTTCTGCCCACCTGCTTT 365
          ||||||||||||||||||||||||||||| |||| |||||||||||||||||||||||||||||
Db  52265 TCAGCGGAACAAAAAGCCTTCTAGGTCTTTTCTTTACCACAAAACAACTTCTTCTGCCCACCTGCTTT 52206

Qy    366 GAAAGGGGCAGAAGTATAGTGGGCGGA-GCTGCCCACCTGCTACAGTGAACAGGGGANCTGGAG 424
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||  ||||||
Db  52205 GAAAGGGGCAGAAGTATAGTGGGCGGAGGCTGCCCACCTGCTACAGTGAACAGGGGGGAG 52146

Qy    425 AAATAACTCACACTTTGGAGGTGGCTTCGCCCCTCTTCATCAGCCGAGCTCTCAACTTAAGCCAATG 484
          |||||||||||||||| |||||| ||||||| |||||||||||||||||||||||||||||||||
Db  52145 AAATAACTCACACTTTGAGGTGCCTCGCCCTTCATCAGCCGAGCTCTCAACTTAAGCCAATG 52086

Qy    485 ACCCCACGGGGAG-CTTACACAAGTTCAAACAGGCCCAAAATGCATTCAATGGACGAGGGGGGAG 543
          |||||||||| || |||||||||| |||||||||||||||||||||||||||||||||||||
Db  52085 ACCCCACGGGAGAGCCTTACACAAGTCAAACAGGCCCAAAATGCATTCAATGGACGAGGGGGAG 52026

Qy    544 GCCAAAGGACTCCGGAGGAGAGAGGGCCCAATAAGGCTGGTTCGATCCAATAGAG 603
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  52025 GCCAAAGGACTCCGGAGGAGAGGGCCCAATAAGGCTGGTTCGATCCAATAGAG 51966

Qy    604 AGAGCAGAGGTGGGCGGAGGGCCCTTTGATTAAATGATCATTCTTGAAAGCAAGCTTCAAAA 663
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  51965 AGAGCAGAGGTGGGCGGAGGGCCCTTTGATTAAATGATCATTCTTGAAAGCAAGCTTCAAAA 51906

Qy    664 TCCGGGTATGCCCGGGTGAGAAATGAGCGACTAAACACCTGGGTGTCATGGCAAGCTTCCA 723
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  51905 TCCGGGTATGCCCGGGTGAGAAATGAGCGACTAAACACCTGGGTGTCATGGCAAGCTCCA 51846

Qy    724 GGGCCGACTGGCCAGAGACAGAGATCCCCAGAGAGGGCCTTGCGAGCCCAGCTCTTCGTGGCCAAGCC 783
          ||||||||||||||||||| |||||||||||||| |||||||||||||||||||||||||||||||
Db  51845 GGGCCGACTGGCCAGAGACAGATCCCCAGAGAGGCCTTGCGAGCCCAGCTCTTCGTGGCCAAGCC 51786

Qy    784 ACTCGGATTTGAACCCCGGGTCCTCAAAGTCAGCTGTGTAGCCTTGAATGAAATCACGTGG 843
          |||||||||||||||||||| || |||| ||  ||| ||||||||| ||||||||||| ||
Db  51785 ACTCGGATTTGAACCCCGGGCTCTTCAAGGTCGTGTTAGCCTTGAAATGAAATCACCTGC 51726

Qy    844 TATGACCAAT 853
          ||||||||||
Db  51725 TATGACCAAT 51716

RESULT 5
ADC21010/c
ID   ADC21010   standard; DNA; 58181 BP.
XX
AC   ADC21010;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human secreted protein-related DNA sequence #428.
XX
KW   gene therapy; human; secreted protein; haemopoietic disorder;
KW   haematological disorder; anaemia; haemophilia; inflammatory disorder;
KW   inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
KW   leukaemia; wound healing; epithelial cell proliferation disorder;
KW   immune disorder; autoimmune disorder; asthmatic disorder;
KW   cardiovascular disorder; atherosclerosis; myocarditis;
KW   infectious disease; HIV; AIDS; endocrine disorder; diabetes;
KW   gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO20092787-A2.
XX
PD   21-NOV-2002.
XX
PF   26-MAR-2002; 2002WO-US009257.
XX
PR   27-MAR-2001; 2001US-0278650P.
PR   12-SEP-2001; 2001US-00950082.
PR   12-SEP-2001; 2001US-00950083.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,    Ruben SM;
XX
DR   WPI; 2003-129287/12.
XX
PT   New human secreted proteins and nucleic acid molecules, useful for
```

```
PT   preparing a diagnostic or pharmaceutical composition for diagnosing,
PT   preventing or treating hematopoietic or hematologic disorders, e.g.
PT   anemia or hemophilia.
XX
PS   Disclosure; SEQ ID NO 964; 1512pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of human
CC   secreted proteins. The DNA and protein sequences of the invention are
CC   useful for detecting, preventing, diagnosing, prognosticating, treating
CC   or ameliorating: haematopoietic or haematological disorders (e.g. anaemia
CC   and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
CC   and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
CC   wound healing and disorders of epithelial cell proliferation; immune
CC   disorders (e.g. autoimmune disorders and asthmatic disorders);
CC   cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC   infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC   and gastrointestinal disorders (e.g. duodenal ulcers and
CC   gastroenteritis). The present DNA sequence was used in the
CC   exemplification of the invention.
XX
SQ   Sequence 58181 BP; 15503 A; 12755 C; 13504 G; 16419 T; 0 U; 0 Other;
```

Query Match   94.9%;   Score 822.6;   DB 9;   Length 58181;
Best Local Similarity   99.3%;   Pred. No. 6.2e-234;
Matches 844;   Conservative 2;   Mismatches 2;   Indels 2;   Gaps 2;

```
Qy      6 TTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAIGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  52565 TTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGCA 52506

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCA 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52505 CGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAAGTGCTGGGATTACAGGCA 52446

Qy    126 TGAGCCACTCGGCCCGCCAGCCGGGTCTCTTTTTAAACATTCCCAGGACTGTACAGCCAACCCAT 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52445 TGAGCCACTCGGCCCGCCAGCCGGGTCTTTTTAAACATTCCCAGGACTGTACAGCCAACCCAT 52386

Qy    186 ACTCACCTGACATTTGGGAACTCCCCCCCCACGGCCCATAACTGATCTGCAGAGGTAAGACC 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52385 ACTCACCTGACATTTGGGAACTCCCCCCCACGGCCCATAACTGATCTGCAGAGGTAAGACC 52326

Qy    246 AAGAGCAAGAAATGGGGGATTCACATCTAAGGTCTTCGTGATGATGGCTGATGAGGAAGAAGRA 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52325 AAGAGCAAGAAATGGGGGATTCACATCTAAGGTCTTCGTGATGATGGCTGATGAGGAAGAAGRA 52266

Qy    306 TCACGGAACAAAAAGCCTCTAGGTCTTCTTTACCACAAACACCTCTCCTGCCCACCTGCTTTT 365
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52265 TCACGGAACAAAAAGCCTCTAGGTCTTCTTTACCACAAACACCTCTCCTGCCCACCTGCTTT 52206

Qy    366 GAAAGGGGCAGAAGTATAGTRGGGGGCA-GCTGCCCACCTGCTACAGTGAAGGGGATCTGGAG 424
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  52205 GAAAGGGGCAGAAGTATAGTRGGGGCGACTGCCCACCTGCTACAGTGAAGGGGATCTGGAG 52146

Qy    425 AAATACTCACACTTGAGGTGCTCGGCTCTCCGCCCTCTTCATCAGCCAGCTCTAACTTAAGCCAATG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52145 AAATACTCACACTTGAGGTGCTCGGCTCTCCGCCCTCTTCATCAGCCAGCTCTAACTTAAGCCAATG 52086

Qy    485 ACCCCACGGGAG-CTTACACAAGTYCAAACAGGGCCCAAAATGCAATTCAATCAITGAGCAGGGGGAG 543
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  52085 ACCCCACGGGAGCTTACACAAGTYCAAACAGGGCCCAAAATGCAATTCAATCAITGAGCAGGGGGAG 52026

Qy    544 GCCAAAGGGACTCCCGGAGGAGAGGGCCCAAATAAGGCTGGTCCTATTTCCGATCCATAGAG 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  52025 GCCAAAGGGACTCCCGGAGGAGAGGGCCCAATAAGGCTGGTCCTATTTCCGATCCATAGAG 51966

Qy    604 AGAGCAAGAGAGTGGGCAGGCCCCCTTTGATTAAATGTAATCATTCTTGAATGCAAGCTTCAAAA 663
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  51965 AGAGCAAGAGAGTGGGCAGGCCCCCTTTGATTAAATGTAATCATTCTTGAATGCAAGCTTCAAAA 51906

Qy    664 TCCGGGTATGCCGGGGGTGAGAATGACAGGACTAACACCTGGGTGTGTCATGGCAAGCCTCCA 723
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  51905 TCCGGGTATGCCGGGGGTGAGAATGACAGGACTAACACCTGGGTGTGTCATGGCAAGCCTCCA 51846
```

```
Qy    724 GGGCCGCAGCTGGCCCAGAGACAGCAGATCCGGAGCAAGAAGAGGCTCTCCAGCAGCTGGTGCCAAGCC 783
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  51845 GGGCCGCAGCTGGCCCAGAGACAGCAGATCCGGAGCAAGAAGAGGCTCTCCAGCAGCTGGTGCCAAGCC 51786

Qy    784 ACTCGGGATTTGAACCCCGGGCTCCTCAAGGTCAGCTTGGGTAGCCTTGANTGAAYCACCTGC 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  51785 ACTCGGGATTTGAACCCCGGGCTCCTTCAAGGTCAGCTGTAGCCTTGAATGAATCACCTGC 51726

Qy    844 TATGACCAAT 853
          ||||||||||
Db  51725 TATGACCAAT 51716
```

```
RESULT 6
ADC86430
XX
DE   Human GPCR gene SEQ ID NO:883.
XX
ID   ADC86430 standard; DNA; 39703 BP.
XX
AC   ADC86430;
XX
DT   01-JAN-2004 (first entry)
XX
PD   02-JAN-2003.
XX
PN   EP1270724-A2.
XX
PF   18-JUN-2002; 2002EP-00013517.
XX
PR   18-JUN-2001; 2001JP-00246789.
XX
PA   (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Suwa M, Asai K, Akiyama Y, Aburatani H;
XX
DR   WPI; 2003-315783/31.
DR   P-PSDB; ADC86431.
XX
PT   New polynucleotide, useful for preparing a composition for treating a
PT   patient in need of increased or suppressed activity or expression of the
PT   guanosine triphosphate-binding protein coupled receptor.
XX
PS   Claim 1; SEQ ID NO 883; 28pp; English.
XX
CC   The invention relates to a novel polynucleotide encoding a guanosine
CC   triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC   the invention may have a use in gene therapy. The polynucleotide and
CC   polypeptide are useful for preparing a composition for treating a patient
CC   in need of increased or suppressed activity or expression of the
CC   guanosine triphosphate-binding protein coupled receptor. The
CC   polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC   invention.
XX
SQ   Sequence 39703 BP; 10210 A; 8488 C; 9860 G; 11145 T; 0 U; 0 Other;
```

Query Match   15.4%;   Score 133.6;   DB 9;   Length 39703;
Best Local Similarity   93.9%;   Pred. No. 1.5e-28;
Matches 139;   Conservative 0;   Mismatches 9;   Indels 0;   Gaps 0;

```
Qy      6 TTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAIGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
          ||| |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  10987 TTTTTTTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 11046

Qy     66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCA 125
          |||  ||||||||||||||||||||||||||||||||||||||||||  |||||||||||||
Db  11047 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAAATGCTGGGATTACAGGCG 11106
```

Qy   126 TGAGGCCACTGCGCCCAGCCGCGGTCTTTTT 153
          |||||||| |||||||| ||||||| | ||||||
Db 11107 TGAGGCCACCGCGCGCCCAGCCTTTTTTTTTT 11134

RESULT 7
ACA64895/c
ID   ACA64895 standard; DNA; 134292 BP.
XX
AC   ACA64895;
XX
DT   27-JUN-2003 (first entry)
XX
DE   Human GABBR1 DNA corresponding to AL031983.
XX
KW   Human; chronic inflammatory joint disease; infection; tumour;
KW   antiinflammatory; cytostatic; antiarthritic; antirheumatic;
KW   immunosuppressive; gene therapy; etiological pathogenicity; ds.
XX
OS   Homo sapiens.
XX
PN   DE10127572-A1.
XX
PD   05-DEC-2002.
XX
PF   30-MAY-2001; 2001DE-01027572.
XX
PR   30-MAY-2001; 2001DE-01027572.
XX
PA   (PATH-) PATHOARRAY GMBH.
XX
PI   Haeupl T, Ungethuem U, Blaess S;
XX
DR   WPI; 2003-240797/24.
XX
PT   Reagents for diagnosis, study and therapy of chronic inflammatory joint
PT   and other diseases, comprises any of many specified genes or derived
PT   proteins.
XX
PS   Claim 1; Page; 12pp; German.
XX
CC   This invention describes a novel reagent for diagnosis, molecular
CC   definition and therapy of chronic inflammatory joint diseases, and other
CC   inflammatory disorders, infective or tumour diseases in humans. The
CC   products of the invention have antiinflammatory, cytostatic,
CC   antiarthritic, antirheumatic and immunosuppressive activity and can be
CC   used for gene therapy. The reagent of the invention and any proteins and
CC   antibodies derived from it, are used (i) for analysing tissue and blood
CC   samples for medical diagnosis; (ii) for diagnosis and characterisation of
CC   chronic joint diseases, on the basis of molecular characterisation, and
CC   determining the etiological pathogenicity principle of as yet
CC   uncharacterised inflammatory diseases, also monitoring progression and/or
CC   treatment of disease, and optimisation of therapy and (iii) for
CC   developing treatments for inflammatory diseases, particularly of joints,
CC   infections and tumours. ACA64801-ACA64965 represent human polynucleotides
CC   used in the method of the invention
XX
SQ   Sequence 134292 BP; 38020 A; 31344 C; 29670 G; 35258 T; 0 U; 0 Other;

Query Match   15.4%; Score 133.6; DB 7; Length 134292;
Best Local Similarity   93.9%; Pred. No. 2.5e-28;
Matches 139; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTTCACCGTGGTTAGCCAGGATGGTCT 65
          ||||||||||||||||||||| ||||||||||||||||| |||||||||||||||||||||||
Db 113399 TTTTTTTTTTTTTTTTTTTTTTAGTAGGAGACGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 113340

Qy   66 CGAACTTCCTGACCTCGTGATCCGCGCCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCA 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 113339 CGATCTCCTGACCTCGTGATCCGCGCCCTCCGGCTTCCCAAAATGCTGGGATTACAGGGCG 113280

Qy   126 TGAGGCCACTGCGCCCAGCCGCGGTCTTTTT 153

Db 113279 TGAGGCCACCGCGCGCCCAGCCTTTTTTTTTTT | |||||| 113252
           |||||||| |||||| ||||||||| | |||||

RESULT 8
AAH26497
ID   AAH26497 standard; DNA; 22255 BP.
XX
AC   AAH26497;
XX
DT   12-NOV-2001 (first entry)
XX
DE   Human low density lipoprotein binding protein 3 (LBP-3) gene.
XX
KW   Low density lipoprotein binding protein 3; LBP-3; LDL; human;
KW   atherosclerosis; antiaterosclerotic; gene therapy; diagnosis; vaccine;
KW   ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   2100..16932
FT   /*tag= a
FT   /note= "contains introns"
FT   exon   2100..2268
FT   /*tag= b
FT   intron   2269..2996
FT   /*tag= c
FT   exon   2997..3332
FT   /*tag= d
FT   intron   3333..6309
FT   /*tag= e
FT   exon   6310..6401
FT   /*tag= f
FT   intron   6402..9738
FT   /*tag= g
FT   exon   9739..9909
FT   /*tag= h
FT   intron   9910..11839
FT   /*tag= i
FT   exon   11840..12034
FT   /*tag= j
FT   intron   12035..14095
FT   /*tag= k
FT   exon   14096..14215
FT   /*tag= l
FT   intron   14216..14443
FT   /*tag= m
FT   exon   14444..14518
FT   /*tag= n
FT   intron   14519..14944
FT   /*tag= o
FT   exon   14945..15037
FT   /*tag= p
FT   intron   15038..15801
FT   /*tag= q
FT   exon   15802..15897
FT   /*tag= r
FT   intron   15898..16638
FT   /*tag= s
FT   exon   16639..16932
FT   /*tag= t
XX
PN   WO200164874-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US005356.
XX
PR   02-MAR-2000; 2000US-00517849.
PR   14-JUL-2000; 2000US-00616289.
XX
PA   (BOST-) BOSTON HEART FOUND INC.

```
XX
PI   Lees AM, Lees RS, Law SW, Arjona AA;
XX
DR   WPI; 2001-565505/63.
DR   P-PSDB; AAB82809.
XX
PT   New isolated low density lipoprotein binding polypeptide for treating,
PT   diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX
PS   Example 4; Fig 24; 143pp; English.
XX
CC   The present sequence is that of genomic DNA encoding novel human low
CC   density lipoprotein binding protein 3 (LBP-3, see AAB82809). The DNA was
CC   isolated from a human genomic library by screening with LBP-3 cDNA (see
CC   AAH26501). The open reading frame spans 10 exons. Human LBP-2 nucleic
CC   acids are among claimed polynucleotides of the invention that encode
CC   novel polypeptides, termed LBPs, capable of binding to native and
CC   methylated LDL. Also claimed are isolated LBP polypeptides and
CC   biologically active fragments and analogues of them, as well as
CC   expression vectors, cells and methods of producing the LBPs. Methods of
CC   determining if an animal is at risk for atherosclerosis, and methods for
CC   evaluating an agent for use in treating atherosclerosis, methods for
CC   treating a cell having an abnormality in structure or metabolism of LBP
CC   are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
CC   nucleic acid, and vaccine compositions, are also claimed
SQ   Sequence 22255 BP; 5195 A; 5302 C; 5924 G; 5834 T; 0 U; 0 Other;

Query Match     15.0%;  Score 130.4;  DB 5;  Length 22255;
Best Local Similarity    89.7%;  Pred. No. 1.1e-27;
Matches  140;  Conservative   0;  Mismatches  16;  Indels   0;  Gaps   0;

Qy          6 TTTTTTTTTTTTTTTTTTTTAAGTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              ||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db       7346 TTTTTTTTTTTTTTTTTTTAGTAGAGACCGGGGTTTCACCGTGTTAGCCAGGATGGTCT 7405

Qy         66 CGAACTCCTGACCTCGTGATCCCGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCA 125
              |||| ||| ||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db       7406 CGATCTCTTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCT 7465

Qy        126 TGAGGCCACTGCGCCCAGCCCGGTCCTTATTTATTTATTTAAACATTC 161
              |||| |||| ||| ||||||||   ||||||||||||||| |||
Db       7466 TGAGCCACCACGCCGCTGGGCCTTATTTATTTATTTATTC 7501

RESULT 9
ID   ACD13448 standard; DNA; 115756 BP.
AC   ACD13448;
DT   13-AUG-2003 (first entry)
DE   Human DNA encoding a p53 modifier, SEQ ID 68.
KW   Human; ds; gene; p53 modifier; cytostatic; cancer; cytostatic;
KW   antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
KW   kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW   apoptotic disorder; cell proliferation disorder.
OS   Homo sapiens.
XX
PN   WO20029122-A1.
XX
PD   12-DEC-2002.
XX
PF   03-JUN-2002; 2002WO-US017382.
XX
PR   05-JUN-2001; 2001US-0296076P.
PR   10-OCT-2001; 2001US-0328605P.
PR   15-FEB-2002; 2002US-0357253P.
XX
PA   (EXEL-) EXELIXIS INC.
```

```
XX
PI   Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR   WPI; 2003-156859/15.
DR   P-PSDB; ABO07223.
XX
PT   Identifying modulators of the p53 pathway for use in treating apoptotic
PT   or cell proliferation disorders, comprises screening for agents that
PT   modulate activity of a human ortholog of genes that modify the p53
PT   pathway in Drosophila.
XX
PS   Example 2; Page 320-351; 678pp; English.
XX
CC   The invention relates to identifying modulators (M1) a candidate p53 pathway
CC   modulating agent, by contacting an assay system comprising a purified HM
CC   polypeptide (human orthologue of genes that modify the p53 pathway in
CC   Drosophila) or nucleic acid with a test agent under conditions, where but
CC   for the presence of the test agent, the system provides a reference
CC   activity, and detecting a test agent-biased activity of the assay system.
CC   Also included are modulating (M2) a p53 pathway of a cell (comprising
CC   contacting a cell defective in p53 function with a candidate modulator
CC   that specifically binds to a HM polypeptide comprising an HM amino acid
CC   sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC   in a mammalian cell (comprising contacting the cell with an agent that
CC   specifically binds an HM polypeptide or nucleic acid) and diagnosing
CC   a disease in a patient (comprising: (a) obtaining a biological sample
CC   from the patient; (b) contacting the sample with a probe for HM
CC   expression; (c) comparing the results with a control; and (d) determining
CC   whether the comparison indicates a likelihood disease). (M1) is useful
CC   for identifying modulators of the p53 pathway. A probe for HM expression
CC   is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC   in a patient, where the cancer has greater than 25 % expression level.
CC   Modulators identified by (M) are useful in a variety of diagnostic and
CC   therapeutic applications, where disease or disorder prognosis is related
CC   to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC   proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC   M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC   the p53 function of the cell, so that the cell undergoes normal
CC   proliferation or progression through the cell cycle. (M2) and (M3) are
CC   also useful for treating defects in the p53 pathway such as angiogenic,
CC   apoptotic or cell proliferation disorders. The present sequence is an HM
CC   nucleic acid encoding a p53 pathway modifying protein
SQ   Sequence 115756 BP; 29575 A; 26665 C; 27977 G; 31539 T; 0 U; 0 Other;

Query Match     15.0%;  Score 130.4;  DB 7;  Length 115756;
Best Local Similarity    89.7%;  Pred. No. 2.1e-27;
Matches  140;  Conservative   0;  Mismatches  16;  Indels   0;  Gaps   0;

Qy          6 TTTTTTTTTTTTTTTTTTTTTAAGTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              ||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db      74889 TTTTTTTTTTTTTTTTTTTAGTAGAGACCGGGGTTTCACCGTGTTAGCCAGGATGGTCT 74948

Qy         66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCGCCTCCCAAAGTGCTGGGATTACAGGCA 125
              |||| ||| |||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db      74949 CGATCTCTTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCT 75008

Qy        126 TGAGGCCACTGCGCCCAGCCCGGTCCTTATTTATTTATTTAAACATTC 161
              |||| |||| ||| ||||||||   ||||||||||||||| |||
Db      75009 TGAGCCACCACGCCGCTGGGCCTTATTTATTTATTTATTC 75044

RESULT 10
ID   AAL53468 standard; DNA; 20951 BP.
AC   AAL53468;
DT   16-JAN-2003 (first entry)
DE   Genomic DNA encoding human Ras-like protein.
KW   Human; Anti-HIV; neuroprotective; nootropic; cerebroprotective; cardiant;
```

```
KW   ophthalmological; antiinflammatory; antianaemic; vasotropic;
KW   immunomodulatory; viruicide; cytostatic; osteopathic; Ras inhibitor; AIDS;
KW   gene therapy; human Ras-like protein; cell proliferation; apoptosis;
KW   genetic immunodeficiency; neurodegenerative disease; Alzheimer's disease;
KW   Parkinson's disease; amyotrophic lateral sclerosis; retinitis pigmentosa;
KW   cerebellar degeneration; aplastic anaemia; ischemic injury; cirrhosis;
KW   viral infection; myocardial infarction; alcohol-induced liver damage;
KW   stroke; toxin-induced disease; lathyrism; cachexia; hepatitis B; cancer;
KW   hepatitis C; osteoporosis; transgenic animal; hepatotropic; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             3000..18568
FT                   /*tag= a
FT                   /product= "Human Ras-like protein"
FT                   /note= "Coding sequence contains 14 introns"
FT   exon            3000..3072
FT                   /*tag= b
FT                   /number= 1
FT   intron          3073..3753
FT                   /*tag= c
FT                   /number= 1
FT   exon            3754..3855
FT                   /*tag= d
FT                   /number= 2
FT   intron          3856..4363
FT                   /*tag= e
FT                   /number= 2
FT   exon            4364..4427
FT                   /*tag= f
FT                   /number= 3
FT   intron          4428..4786
FT                   /*tag= g
FT                   /number= 3
FT   exon            4787..4918
FT                   /*tag= h
FT                   /number= 4
FT   intron          4919..5702
FT                   /*tag= i
FT                   /number= 4
FT   exon            5703..5853
FT                   /*tag= j
FT                   /number= 5
FT   intron          5854..6056
FT                   /*tag= k
FT                   /number= 5
FT   variation       replace(5539,G)
FT                   /*tag= ae
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(5658,G)
FT                   /*tag= af
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(5861,T)
FT                   /*tag= ag
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(6023,G)
FT                   /*tag= ah
FT                   /standard_name= "Single nucleotide polymorphism"
FT   exon            6057..6230
FT                   /*tag= l
FT                   /number= 6
FT   intron          6231..6389
FT                   /*tag= m
FT                   /number= 6
FT   exon            6390..6506
FT                   /*tag= n
FT                   /number= 7
FT   intron          6507..8832
FT                   /*tag= o
FT                   /number= 7
FT   variation       replace(6799,T)
FT                   /*tag= ai
FT                   /standard_name= "Single nucleotide polymorphism"

FT   exon            8833..9114
FT                   /*tag= p
FT                   /number= 8
FT                   /standard_name= "Single nucleotide polymorphism"
FT   intron          9115..9885
FT                   /*tag= q
FT                   /number= 8
FT   exon            9886..9963
FT                   /*tag= r
FT                   /number= 9
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(9842,C)
FT                   /*tag= ak
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(9579,A)
FT                   /*tag= aj
FT                   /standard_name= "Single nucleotide polymorphism"
FT   intron          9964..10201
FT                   /*tag= s
FT                   /number= 9
FT   exon            10202..10324
FT                   /*tag= t
FT                   /number= 10
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(10159,C)
FT                   /*tag= al
FT   intron          10325..10638
FT                   /*tag= u
FT                   /number= 10
FT   exon            10639..10754
FT                   /*tag= v
FT                   /number= 11
FT   intron          10755..15675
FT                   /*tag= w
FT                   /number= 11
FT   variation       replace(12025,G)
FT                   /*tag= am
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(14723,C)
FT                   /*tag= an
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(14996,A)
FT                   /*tag= ao
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       15676..15817
FT                   /*tag= 
FT                   /standard_name= "Single nucleotide polymorphism"
FT   exon            15818..16071
FT                   /*tag= x
FT                   /number= 12
FT   intron          16072..16108
FT                   /*tag= y
FT                   /number= 12
FT   exon            16109..16828
FT                   /*tag= z
FT                   /number= 13
FT   variation       replace(16153,G)
FT                   /*tag= ap
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(16181,A)
FT                   /*tag= aq
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(16756,G)
FT                   /*tag= ar
FT                   /standard_name= "Single nucleotide polymorphism"
FT   exon            16829..17008
FT                   /*tag= ab
FT                   /number= 14
FT   intron          17009..18491
FT                   /*tag= ac
FT                   /number= 14
FT   variation       replace(18059,G)
FT                   /*tag= as
FT                   /standard_name= "Single nucleotide polymorphism"
```

```
FT   variation        replace(18364,T)
FT                    /*tag= at
FT                    /standard_name= "Single nucleotide polymorphism"
FT   exon             18492..18568
FT                    /*tag= ad
FT                    /number= 15
FT   variation        replace(18861,A)
FT                    /*tag= au
FT                    /standard_name= "Single nucleotide polymorphism"
FT   variation        replace(20443,A)
FT                    /*tag= av
FT                    /standard_name= "Single nucleotide polymorphism"
FT   variation        replace(20881,T)
FT                    /*tag= aw
FT                    /standard_name= "Single nucleotide polymorphism"
XX
PN   WO020272765-A2.
XX
PD   19-SEP-2002.
XX
PF   08-MAR-2002; 2002WO-US007159.
XX
PR   14-MAR-2001; 2001US-0805455.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Gan W, Ye J, Di Francesco V, Beasley EM;
XX
DR   WPI; 2002-750490/81.
DR   P-PSDB; AAO26354.
XX
PT   New isolated Ras-like protein polypeptides, useful for treating AIDS,
PT   neurodegenerative diseases, ischemic injuries, toxin-induced diseases,
PT   viral infections, cancer and osteoporosis.
XX
PS   Claim 4; Fig 3A-G; 91pp; English.
XX
CC   The invention relates to a novel human Ras-like protein comprising a
CC   fully defined sequence of 615 amino acids, given in the specification,
CC   its allelic variant, orthologue or its fragment comprising at least 10
CC   contiguous amino acids. The new human Ras-like proteins and the
CC   polypeptides encoding them are useful in the diagnosis, prevention, and
CC   treatment of inflammation and disorders associated with cell
CC   proliferation and apoptosis. They are also useful in treating AIDS and
CC   other infectious or genetic immunodeficiencies, neurodegenerative
```

Query Match   14.9%; Score 129.4; DB 6; Length 20951;
Best Local Similarity   87.1%; Pred. No. 2e-27;
Matches 142; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy       7 TTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGGAGATGGGGGTTTCACCGTGTTAGCCAGGATG 66
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17276 TTTTTTTTTTTTGTATTTTTAGTAAAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTC 17217

Qy      67 GAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCAT 126
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17216 GATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGT 17157

Qy     127 GAGCCACTGCGCCCAGCCGGGTCTTTTTAAACATTCCCAGGAC 169
           ||||||||||||||||||||||||||
Db   17156 GAGCCACCGCGCCCGGCCGCTGTCCCTTCTTCTTTATTATCCACCAC 17114
```

```
RESULT 11
ADE11169_3
Continuation (4 of 4) of ADE11169 from base 300001 (Human transporter protein encoding
WP   Sequence split into 4 fragments LOCUS ADE11169 Accession Ade11169
WP   Fragment Name           Begin        End
WP   ADE11169_0              1            110000
WP   ADE11169_1              100001       210000
WP   ADE11169_2              200001       310000
WP   ADE11169_3              300001       394191
```

```
RESULT 12
ABV12806/c
ID   ABV12806 standard; cDNA; 407 BP.
XX
AC   ABV12806;
XX
DT   13-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 12797.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO020160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0199862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R, Endege WO, Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 2109; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 407 BP; 124 A; 111 C; 81 G; 91 T; 0 U; 0 Other;
```

Query Match   14.9%; Score 128.8; DB 9; Length 94191;
Best Local Similarity   86.6%; Pred. No. 5.8e-27;
Matches 142; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy       2 CGAGTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGGATGGGGGTTTCACCGTGTTAGCCAGGATG 61
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   93119 CTRAPTTTTTTTTTTTTTTTTTTTTATTAGRACANGGGTTTCACCGTGTTAGCCAGGATG 93178

Qy      62 GTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACA 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   93179 ATCTCGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACA 93238

Qy     122 GGCATGAGCCACTGCGCCCACGCGCGGTCTCTTTTTAAACATTCCCA 165
           |||||||||        ||||              |||||||||
Db   93239 GGCATGAGCGTACTCGCGCCCCTGCCGCCCAAAATTGTAATCTTTAAACA 93282
```

Query Match   14.8%; Score 128; DB 5; Length 407;

RESULT 13

```
ID   ABV03637/c
QY   Best Local Similarity   93.1%;
     Matches 134;  Conservative   0;  Pred. No.  1e-27;  Indels   0;  Gaps   0;
                                      Mismatches 10;

Qy             6 TTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTCACCGTGTTAGCCAGGATGGTCT 65
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           389 TTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTCACCGTGTTAGCCAGGATGGTCT 369

Qy            66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCCA 125
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           329 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 270

Qy           126 TGAGCCACTGCGCCCAGCCGGGTCT 149
                 ||||||||||||||||||||| |||
Db           269 TGAGCCACCGCGCCCGGGCCTATTT 246
```

```
ID   ABV03637;
XX
AC
XX
DT   13-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 3628.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R, Endege WO, Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 653; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 447 BP; 126 A; 121 C; 100 G; 97 T; 0 U; 3 Other;
```

Query Match   14.8%;   Score 128; DB 5; Length 447;
Best Local Similarity 93.1%;   Pred. No. 1e-27;

RESULT 14

```
ID   ABV33935/c
QY   Matches 134;  Conservative   0;  Mismatches 10;  Indels   0;  Gaps   0;
                                 Pred. No.  1.1e-27;

Qy             6 TTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTCACCGTGTTAGCCAGGATGGTCT 65
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           428 TTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTCACCGTGTTAGCCAGGATGGTCT 369

Qy            66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCCA 125
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           368 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 309

Qy           126 TGAGCCACTGCGCCCAGCCGGGTCT 149
                 ||||||||||||||||||||| |||
Db           308 TGAGCCACCGCGCCCGGGCCTATTT 285
```

```
ID   ABV33935;
XX
AC
XX
DT   16-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 33926.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacogenomic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R, Endege WO, Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 7156; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (II) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 458 BP; 135 A; 125 C; 99 G; 99 T; 0 U; 0 Other;
```

Query Match   14.8%;   Score 128; DB 5; Length 458;
Best Local Similarity 93.1%;   Pred. No. 1.1e-27;
Matches 134; Conservative   0; Mismatches 10; Indels   0; Gaps   0;

```
Qy       6 TTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     440 TTTTTTTTTTTTTTTTTTTTTTTTTTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 381

Qy      66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     380 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 321

Qy     126 TGAGCCACTGCGCCCAGCCGGTCT 149
           |||||||| ||||||| ||| ||||
Db     320 TGAGCCACCGCGGCCCCGGCCCGGCCTATTT 297
```

RESULT 15

ABV42808/c

ID   ABV42808 standard; cDNA; 458 BP.
XX
AC   ABV42808;
XX
DT   16-SEP-2002 (first entry)
XX
DE   Human prostate expression marker cDNA 42799.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-02074545P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R, Endege WO, Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 8557-8558; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement.(II) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 458 BP; 135 A; 125 C; 99 G; 99 T; 0 U; 0 Other;

Query Match   14.8%; Score 128; DB 5; Length 458;
Best Local Similarity   93.1%; Pred. No. 1.1e-27;
Matches 134; Conservative   0; Mismatches   10; Indels   0; Gaps   0;

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 24, 2004, 13:21:26 ; Search time 2500 Seconds
                     (without alignments)
                     15031.364 Million cell updates/sec

Title: US-10-621-363-12
Perfect score: 867
Sequence: 1 tcgagtttttttttttttt............accaatctcgtgccgaattc 867

Scoring table: IDENTITY_NUC
                     Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151699S residues

Total number of hits satisfying chosen parameters:

Post-processing: Minimum Match 0%
                     Maximum Match 100%
                     Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   GenEmbl:*
    1:   gb_ba:*
    2:   gb_htg:*
    3:   gb_in:*
    4:   gb_om:*
    5:   gb_ov:*
    6:   gb_pat:*
    7:   gb_ph:*
    8:   gb_pl:*
    9:   gb_pr:*
   10:   gb_ro:*
   11:   gb_sts:*
   12:   gb_sy:*
   13:   gb_un:*
   14:   gb_vi:*
   15:   em_ba:*
   16:   em_fun:*
   17:   em_hum:*
   18:   em_in:*
   19:   em_mu:*
   20:   em_om:*
   21:   em_or:*
   22:   em_ov:*
   23:   em_pat:*
   24:   em_ph:*
   25:   em_pl:*
   26:   em_ro:*
   27:   em_sts:*
   28:   em_un:*
   29:   em_vi:*
   30:   em_vrt:*
   31:   em_htg_hum:*
   32:   em_htg_inv:*
   33:   em_htg_mus:*
   34:   em_htg_other:*
   35:   em_htg_pln:*
   36:   em_htg_rod:*
   37:   em_htg_mam:*
   38:   em_htg_vrt:*
   39:   em_sy:*
   40:   em_htgo_hum:*
   41:   em_htgo_mus:*
   42:   em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
c 1	822.6	94.9	124874	HS888M10	AL031296 Human DNA
c 2	269	31.0	2069	AK095438	AK095438 Human sapi
3	137.8	15.9	152037	HS167A19	AL031427 Human DNA
4	133.6	15.4	39703	AX646691	AX646691 Sequence
5	133.6	15.4	39703	AB054476	AB054476 Homo sapi
6	133.6	15.4	100000	AB000688	AB000688 Human DNA
c 7	133.6	15.4	102684	AP000079	AP000079 Human DNA
c 8	133.6	15.4	129806	AC006137	AC006137 Homo sapi
c 9	133.6	15.4	134292	HS271M21	AL031983 Human DNA
c 10	133.6	15.4	145431	AL662826	AL662826 Human DNA
c 11	133.6	15.4	155586	AP006245	AP006245 Homo sapi
c 12	133.6	15.4	155874	AL645936	AL645936 Human DNA
c 13	133.6	15.4	164872	AC138080	AC138080 Homo sapi
c 14	133.6	15.4	199706	AC097329	AC097329 Pan trogl
15	133.6	15.4	207362	AL355484	AL355484 Homo sapi
c 16	133	15.3	166336	AL445467	AL445467 Human DNA
17	132.8	15.3	175269	AC145354	AC145354 Gorilla g
18	132.6	15.3	181049	AC093131	AC093131 Papio anu
19	132.2	15.2	160099	AC145498	AC145498 Papio anu
20	132.2	15.2	165271	AC011032	AC011032 Homo sapi
21	131.6	15.2	129711	AC110768	AC110768 Homo sapi
c 22	131.6	15.2	172945	AC067719	AC067719 Homo sapi
c 23	131.6	15.2	193306	AC025512	AC025512 Homo sapi
c 24	131.6	15.2	194638	AC112913	AC112913 Homo sapi
25	131.4	15.2	154546	AC093140	AC093140 Pan trogl
26	131.4	15.2	165024	AC093707	AC093707 Pan trogl
c 27	131.4	15.2	178367	AC006115	AC006115 Homo sapi
28	131.4	15.2	284374	AC162240	AC162240 Pan trogl
c 29	131	15.1	91093	AP004551	AP004551 Homo sapi
c 30	131	15.1	158562	AL603888	AL603888 Human DNA
c 31	131	15.1	161696	AC023129	AC023129 Homo sapi
c 32	131	15.1	179335	AC147192	AC147192 Papio anu
33	131	15.1	199522	AC089999	AC089999 Homo sapi
34	131	15.1	203424	AC106884	AC106884 Homo sapi
c 35	131	15.1	205342	AC068282	AC068282 Homo sapi
c 36	130.8	15.1	42588	AC079360	AC079360 Homo sapi
37	130.8	15.1	98692	AL391730	AL391730 Human DNA
38	130.8	15.1	132592	AL589792	AL589792 Human sapi
39	130.8	15.1	150902	AC018391	AC018391 Homo sapi
40	130.8	15.1	166518	HS10G022	AL031847 Human DNA
41	130.8	15.1	170141	AC015479	AC015479 Homo sapi
42	130.8	15.1	239008	AC022460	AC022460 Homo sapi
43	130.6	15.1	109403	AC025916	AC025916 Homo sapi
44	130.6	15.1	168000	AC062030	AC062030 Homo sapi
45	130.6	15.1	171166	AC012354	AC012354 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS   HS888M10   124874 bp   DNA   linear   PRI   04-MAR-2003
DEFINITION   Human DNA sequence from clone RP5-888M10 on chromosome
             1p36.11-36.31, complete sequence.
ACCESSION   AL031296
VERSION   AL031296.1   GI:4106564
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 124874)
AUTHORS   Cobley,V.
TITLE   Direct Submission

JOURNAL   Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
           Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequests@sanger.ac.uk
           humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
           On Jan 5, 1999 this sequence version replaced gi:3881974.

COMMENT   -------------- Genome Center
           Center: Wellcome Trust Sanger Institute
           Center code: SC
           Web site: http://www.sanger.ac.uk
           Contact: humquery@sanger.ac.uk

           During sequence assembly data is compared from overlapping clones.
           Where differences are found these are annotated as variations
           together with a note of the overlapping clone name. Note that the
           variation annotation may not be found in the sequence submission
           corresponding to the overlapping clone, as we submit sequences with
           only a small overlap as described above.
           This sequence was finished as follows unless otherwise noted: all
           regions were either double-stranded or sequenced with an alternate
           chemistry or covered by high quality data (i.e., phred quality >=
           30); an attempt was made to resolve all sequencing problems, such
           as compressions and repeats; all regions were covered by at least
           one plasmid subclone or more than one M13 subclone; and the
           assembly was confirmed by restriction digest, except on the rare
           occasion of the clone being a YAC.
           The following abbreviations are used to associate primary accession
           numbers given in the feature table with their source database:
           Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
           on the WORMPEP database can be found at
           http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
           was generated from part of bacterial clone contigs of human
           chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
           Group. Further information can be found at
           http://www.sanger.ac.uk/HGP/Chr1
           RP5-888M10 is from the library RPCI-5 constructed by the group of
           Pieter de Jong. For further details see
           http://www.chori.org/bacpac/home.htm
           VECTOR: pCYPAC2.

FEATURES   Location/Qualifiers
    source   1..124874
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="RZPD:RPCIP704M10888"
              /db_xref="taxon:9606"
              /chromosome="1"
              /map="p36.11-36.31"
              /clone="RP5-888M10"
              /clone_lib="RPCI-5"

ORIGIN

Query Match   94.9%;   Score 822.6; DB 9;   Length 124874;
Best Local Similarity 99.3%;   Pred. No. 9.7e-240;
Matches 844; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

Qy   6 TTTTTTTTTTTTTTTTTTTTTTAAGTAGAGGATGGGGTTTCACGTGTTAGCCAGGATGGTCTT 65
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 87285 TTTTTTTTTTTTTTTTTTTAAGTAGAGAGGGGGTTTCACGTGTTAGCCAGGATGGTCT 87226

Qy   66 CGAACCCTGACCTCGTGATCCGCCCGCCTGGCCTTCCCAAAGTGTGGGATTACAGGCA 125
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 87225 CGAACTTCCTGACCTCGTGATCCGCCCGCCTTCGGCCTTCCCAAAGTGTGGGATTACAGGCA 87166

Qy   126 TGAGCCACTCCGGCCCAGCCGGGTCTCTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 185
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 87165 TGAGCCACTCGGCCCAGCCGGGTCTCTTTTTAAACATTCCCCAGGACTCTACAGCCAACCCAT 87106

Qy   186 ACTCACCTGACATTTGGGAACATCTCCCCCCCACGGGCCATAACTGATCCTGCACAGGTAAGACC 245
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 87105 ACTCACCTGACATTTGGGAACTCCCCCCCCACGGGCCCCATAACTGATCCTGCACAGGTAAGACC 87046

Qy   246 AAGAGCAAGAGAATGGGGGGATTCACATCTAAGGTCTCGGTGATGGGCTGATGGAAGGAGGAAGAA 305
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 87045 AAGAGCAAGAATGGGGGGATTCACATCTAAGGTCTCGTGATGGCTGATGGAAGGAGGAAGAA 86986

Qy   306 TCAGCGGAACAAAAAGCCTTAGGTCTTTTTCTTACCACAAAACACTTCTTCTGCCCACTTGCTTT 365
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 86985 TCAGCGGAACAAAAAGCCTAGGTCTTTTCTTACCACAAAACAACCTTCTTTCTGCCCACTTGCTTT 86926

Qy   366 GAAAGGGGGCGAGAGTATAGTGGGGGGA-GGTGCCCACCTTGCTACAGTGAAGGGAGTCTGGAG 424
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 86925 GAAAGGGGGCGAGAGTATAGTGGGGGGGCGGTGCCCACCTTGCTACAGTGAAGGGAGTCTGGAG 86866

Qy   425 AAATACTCACACTTTGAGGTGCTGGCTGCCCCTTCATCAGCGACTCTAACTTAAGGCAAATG 484
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 86865 AAATACTCACACTTTGAGGTGCTGGCTGCCCCTTCATCAGCGACTCTAACTTAAGGCAAATG 86806

Qy   485 ACCCCACGGGAG-CTTACACAAGTCAAAACAGGCCCCAAATGCATTCATGAGCAGGGGGAG 543
            |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db 86805 ACCCCACGGGGAGCTTACACAAGTCCAAAACAGGCCCAAATGCATTCATGAGCAGGGGGAG 86746

Qy   544 GCCAAAGGACTCCGGAGGAGGAGGGCCCAATATAAGGCTGGTGCTATTTCCGATCCATAGAG 603
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 86745 GCCAAAGGACTCCGGAGGAGGAGGGCCCAATATAAGGCTGGTGCTATTTCCGATCCATAGAG 86686

Qy   604 AGAGCAGAGAGGTGGGCCAGGCCCCTTTTGATTTAAAGTATCATTCTTGAATGCAAAGCTTCAAAA 663
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 86685 AGAGCAGAGAGGTGGGCCAGGCCCCTTTTGATTTAAAGTATCATTCTTGAATGCAAAGCTTCAAAA 86626

Qy   664 TCCGGGTATGCCGGGGTGAGAATGAGCAGGACTAACACTGGGGTGTCANGGCAAGCCCTCCA 723
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 86625 TCCGGGTATGCCGGGGTGAGAATGAGCAGGACTAACACTGGGGTGTCATGGCAAGCCCTCCA 86566

Qy   724 GGGGCCGACTGCCAGAGACGATCCGCAAGAGGGCTCGCAGCCCAGCCTCGGTGCCCAAGCC 783
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 86565 GGGGCCGACTGCCAGAGACGATCCGCAAGAGGGCTCGCAGCCCAGCCTCGGTGCCCAAGCC 86506

Qy   784 ACTCGGATTTGAACCCCGCCGCTCCCCAAGGTCAGCGTGTGAGCCCTTGAANTGAAAYCACCCTGC 843
            ||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||
Db 86505 ACTCGGATTTGAACCCCGGGCCTCCCCAAGGTCAGCGTGTGAGCCCTTGAATGAAATCACCCTGC 86446

Qy   844 TATGACCAAT 853
            ||||||||||
Db 86445 TATGACCAAT 86436

RESULT 2
AK095438/c

LOCUS   AK095438   2069 bp   mRNA   linear   PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ38119 fis, clone D3OST2004637.
ACCESSION   AK095438
VERSION   AK095438.1 GI:21754694
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
            Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
            Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
            Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
            Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE   NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 2069)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team);
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA full insert
            library

FEATURES   Location/Qualifiers

source   1..2069
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="D3OST2004637"
     /cell_type="CD34+ Cells"
     /clone_lib="D3OST2"
     /note="Cloning vector: pME18SFL3-mRNA from CD34+ cells
     after 3-days ODP induction.-primary culture, CD34+ Cells"

ORIGIN

Query Match   31.0%;   Score 269;   DB 9;   Length 2069;
Best Local Similarity   100.0%;   Pred. No. 1e-70;
Matches 269;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   6   TTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT   65
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   269   TTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT   210

Qy   66   CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCA   125
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   209   CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA   150

Qy   126   TGAGCCACTGCGCCCAGCCGGGTCTTTTTTAAACATTCCCCAGGACTGTACAGCCAACCAT   185
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   149   TGAGCCACTGCGCCCAGCCGGGTCTTTTTTAAACATTCCCCAGGACTGTACAGCCAACCAT   90

Qy   186   ACTCACCTGACATTGGGAACTCCCCCCCCCCAGGCCCCATAACTGATCTCGCAGAGGTTAAGACC   245
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   89   ACTCACCTGACATTGGGAACTCCCCCCCCATAACTGATCTCGCAGAGGTTAAGACC   30

Qy   246   AAGAGCAAGAATGGGGGATTCACATCTAA   274
     |||||||||||||||||||||||||||||
Db   29   AAGAGCAAGAATGGGGGATTCACATCTAA   1

RESULT 3
HS167A19   152037 bp   DNA   linear   PRI 04-MAR-2003
LOCUS   HS167A19   152037 bp   DNA   linear   PRI 04-MAR-2003
DEFINITION   Human DNA sequence from clone RP1-167A19 on chromosome 1p32.1-33,
     complete sequence.
ACCESSION   AL031427
VERSION   AL031427.15   GI:4835258
KEYWORDS   HTG.
SOURCE   Homo sapiens
   ORGANISM   Homo sapiens (human)
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 152037)
AUTHORS   Hall,R.
TITLE   Direct Submission
JOURNAL   Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
     Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
     humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On May 14, 1999 this sequence version replaced gi:4808231.
     ------------------- Genome Center
     Center: Wellcome Trust Sanger Institute
     Center code: SC
     Web site: http://www.sanger.ac.uk
     Contact: humquery@sanger.ac.uk

     During sequence assembly data is compared from overlapping clones.
     Where differences are found these are annotated as variations
     together with a note of the overlapping clone name. Note that the
     variation annotation may not be found in the sequence submission
     corresponding to the overlapping clone, as we submit sequences with
     only a small overlap as described above.

---

construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone, and the
one plasmid subclone or more than one M13 subclone, and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP1-167A19 is from the library RPCI-1 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm

FEATURES   Location/Qualifiers
source   1..152037
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"
     /chromosome="1"
     /map="1p32.1-33"
     /clone="RP1-167A19"
     /clone_lib="RPCI-1"

ORIGIN

Query Match   85.1%;   Score 137.8;   DB 9;   Length 152037;
Best Local Similarity   85.1%;   Pred. No. 1.7e-30;
Matches 154;   Conservative 0;   Mismatches 27;   Indels 0;   Gaps 0;

Qy   6   TTTTTTTTTTTTTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT   65
     ||||||||||||||||||| || |||||||||||||||||||||||||||||||||||||||||||
Db   127313   TTTTTTTTTTTTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT   127372

Qy   66   CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCA   125
     |||||| || ||||| | ||||||||||||||||||||||| |||||||||||||||||||
Db   127373   CAATCTCCTGACCTCATGACCCGCCCGCCTCCGCCTCCCAAAGTGCTGGGATTACAGGCT   127432

Qy   126   TGAGCCACTGCGCCCAGCCGGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCAT   185
     ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   127433   TGAGCCACTGCGCCCAGCCCATATGTTTTTAAAATAAACAACAGTGCTCCACCTCCACAAAA   127492

RESULT 4
AX646691   39703 bp   DNA   linear   PAT 04-MAR-2003
LOCUS   AX646691   39703 bp   DNA   linear   PAT 04-MAR-2003
DEFINITION   Sequence 883 from Patent EP1270724.
ACCESSION   AX646691
VERSION   AX646691.1   GI:28799074
KEYWORDS   .
SOURCE   Homo sapiens
   ORGANISM   Homo sapiens (human)
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.
TITLE   Guanosine triphosphate-binding protein coupled receptors
JOURNAL   Patent: EP 1270724-A 883 02-JAN-2003;
     National Institute of Advanced Industrial Science and Technology
     (JP), Center for Advanced Science and Technology Incubation, Ltd.
     (JP)

FEATURES   Location/Qualifiers

source   1. .39703
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"

CDS   join(201. .524,3968. .4102,4517. .4687,5787. .5888,6700. .6891,
                14506. .14748,15377. .15440,16990. .17067,18612. .18762,
                19258. .19390,20773. .20889,20987. .21167,22295. .22422,
                23003. .23131,23354. .23497,24275. .24444,39365. .39503)
                /codon_start=1
                /note="unnamed protein product"
                /protein_id="CAD69472.1"
                /db_xref="REMTREMBL:CAD69472"
                /translation="MGPGAPFARVGWPLPLLVVMAAGVAPVWASHSPHLPRPHSRVPP
                HPSSERRAVVIGALFPMSGGWPGGQACQPAVEMALEDVNSRRDILPDYELKLIHHDSK
                VALDMGCDPGQATKIYELRLYNDPIKIILMPGCSSVSTLVAEAARMNLIVLSYGSSS
                PALSNRQRPFPFRTHIPSATLHNPTRVKLFEKWGWKKIAIIQQTTEVFTSTLDDLEER
                VKEAGIEITFRQSFSDPAVPVKNLKVYKERLFGKKYVWFLIGWYADNWFKIYDPSIN
                CTVDEMTEAVEGHITTEIVMLNPANTRSISNMTSQEFVEKLTKRLKRHPETGGFQBA
                PLAYDAIWALALALNKTSGGGRSGVRLEDFNYNNQTITDQIYRAMNSSSFEGVSGHV
                VFDASGSRMAWTLIEQLQGGSYKKIGYYDSTKDDLSWSKTDKWIGGSPPADQTLVIKT
                FRFLSQKLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSIAL
                AAVFPLGLDGYHIGRNQFPFVCQARLWLLGLGFSLGYGSMFTKIWWHTVFTKKEBKK
                EWRKILEPWKLYATVGLLVGMDVLTLAIWQIVDPLHRTIEVPLERRCYGQENEQSSVA
                IRALGLCIFYGYKGLLLLGIFLAYETKSVSTEKINDHRAVGMAIYNVAVLCLITAPV
                TMILSSQODAAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSBAAQDTMKTSSST
                NNEEEKSRLLEKENRELELEKIIAEKEERVSELRHQLQSRQQLRSRHPPTPPEPSGGL
                PRGPEPPDRLSCDGSRVHLLYKVNPQHCPGNGSVRDKDEARRHQDEDTGNSHLDLI
                VGGVFTGESN"

ORIGIN

RESULT 5

Query Match   15.4%;   Score 133.6; DB 6; Length 39703;
Best Local Similarity   93.9%;   Pred. No. 3e-29;
Matches 139; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10987 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 11046

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11047 CGATCTCCTGACCTCGTGATCCGCCGCCTCGGCCTCCCAAAATGCTGGGATTACAGGCG 11106

Qy   126 TGAGCCACTGCGCCCCAGCCGGGTCTTTTTT 153
          ||||||||||| |||||| ||||| |||||
Db 11107 TGAGCCACGCGCCCCAGCCTTTTTTTTTT 11134

LOCUS   AB065476   39703 bp   DNA   linear   PRI 23-JUL-2002
DEFINITION Homo sapiens gene for seven transmembrane helix receptor, complete
            cds, isolate:CBRC7TM_39.
ACCESSION   AB065476
VERSION   AB065476.1 GI:21928256
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S.,
            Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
  TITLE   Genome-wide discovery and analysis of human seven transmembrane
            helix receptor genes
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 39703)
  AUTHORS   Suwa,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research
            Center (CBRC), National Institute of Advanced Industrial Science
            and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan
            (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,

---

FEATURES   Location/Qualifiers
     source   1. .39703
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="CBRC7TM_39"
                     /db_xref="taxon:9606"
                     /chromosome="6"
     CDS   join(201. .524,3968. .4102,4517. .4687,5787. .5888,6700. .6891,
                     14506. .14748,15377. .15440,16990. .17067,18612. .18762,
                     19258. .19390,20773. .20889,20987. .21167,22295. .22422,
                     23003. .23131,23354. .23497,24275. .24444,39365. .39503)
                     /codon_start=1
                     /evidence=not_experimental
                     /product="seven transmembrane helix receptor"
                     /protein_id="BAC05310.1"
                     /db_xref="GI:21928257"
                     /translation="MGPGAPFARVGWPLPLLVVMAAGVAPVWASHSPHLPRPHSRVPP
                     HPSSERRAVVIGALFPMSGGWPGGQACQPAVEMALEDVNSRRDILPDYELKLIHHDSK
                     VALDMGCDPGQATKIYELRLYNDPIKIILMPGCSSVSTLVAEAARMNLIVLSYGSSS
                     PALSNRQRPFPFRTHIPSATLHNPTRVKLFEKWGWKKIAIIQQTTEVFTSTLDDLEER
                     VKEAGIEITFRQSFSDPAVPVKNLKVYKERLFGKKYVWFLIGWYADNWFKIYDPSIN
                     CTVDEMTEAVEGHITTEIVMLNPANTRSISNMTSQEFVEKLTKRLKRHPETGGFQBA
                     PLAYDAIWALALALNKTSGGGRSGVRLEDFNYNNQTITDQIYRAMNSSSFEGVSGHV
                     VFDASGSRMAWTLIEQLQGGSYKKIGYYDSTKDDLSWSKTDKWIGGSPPADQTLVIKT
                     FRFLSQKLFISVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNSQPNLNNLTAVGCSIAL
                     AAVFPLGLDGYHIGRNQFPFVCQARLWLLGLGFSLGYGSMFTKIWWHTVFTKKEBKK
                     EWRKILEPWKLYATVGLLVGMDVLTLAIWQIVDPLHRTIEVPLERRCYGQENEQSSVA
                     IRALGLCIFYGYKGLLLLGIFLAYETKSVSTEKINDHRAVGMAIYNVAVLCLITAPV
                     TMILSSQODAAFASLAIVFSSYITLVVLFVPKMRRLITRGEWQSBAAQDTMKTSSST
                     NNEEEKSRLLEKENRELELEKIIAEKEERVSELRHQLQSRQQLRSRHPPTPPEPSGGL
                     PRGPEPPDRLSCDGSRVHLLYKVNPQHCPGNGSVRDKDEARRHQDEDTGNSHLDLI
                     VGGVFTGESN"

ORIGIN

RESULT 6

Query Match   15.4%;   Score 133.6; DB 9; Length 39703;
Best Local Similarity   93.9%;   Pred. No. 3e-29;
Matches 139; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10987 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 11046

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11047 CGATCTCCTGACCTCGTGATCCGCCGCCTCGGCCTCCCAAAATGCTGGGATTACAGGCG 11106

Qy   126 TGAGCCACTGCGCCCCAGCCGGGTCTTTTTT 153
          ||||||||||| |||||| ||||| |||||
Db 11107 TGAGCCACGCGCCCCAGCCTTTTTTTTTT 11134

LOCUS   AP000079   100000 bp   DNA   linear   PRI 07-JAN-2000
DEFINITION Homo sapiens genomic DNA, chromosome 8p11.2, senescence gene
            region, section 15/19, complete sequence.
ACCESSION   AP000079
VERSION   AP000079.1 GI:4580000
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE   1 (sites)
AUTHORS   Isomura,M., Ikegawa,S., Kinjo,T. and Nakamura,Y.
TITLE   DNA sequence analysis of a 1.9-Mb region on chromosome 8p11.2
JOURNAL   Published Only in DataBase (1999)
REFERENCE   2 (bases 1 to 100000)
AUTHORS   Hirakawa,M., Yamaguchi,H., Imai,K. and Shimada,J.
TITLE   Direct Submission
JOURNAL   Submitted (12-FEB-1999) Mika Hirakawa, Japan Science and Technology
             Corporation (JST), Advanced Database Department; 5-3, Yonban-cho,
             Chiyoda-ku, Tokyo 102-0028, Japan (E-mail:mika@tokyo.jst.go.jp,
             Tel:81-3-5214-8491, Fax:81-3-5214-8470)
COMMENT   This sequence is conducted by Japanese sequencing Team.
             Research as a JST sequencing Team.
             Principal Investigator: Yusuke Nakamura Ph.D
             Phone:+81-3-5449-5372, Fax:+81-3-5449-5433,
             yusuke@hgc.ims.u-tokyo.ac.jp
             The sequence is submitted by Human Genome Sequencing in ALIS
             project of JST.
             Japan Science and Technology Corporation (JST)
             5-3, Yonbancyo, Chiyoda-ku, Tokyo, 102-0081 Japan
             For further information about this sequences, please visit our
             sequence archive web site (http://www-alis.tokyo.jst.go.jp/HGS/top.
             html) or send email to webmaster@www-alis.tokyo.jst.go.jp.
FEATURES   Location/Qualifiers
     source   1..100000
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p11.2"
ORIGIN

Query Match   15.4%; Score 133.6; DB 9; Length 100000;
Best Local Similarity   76.9%; Pred. No. 3.2e-29;
Matches 163; Conservative   0; Mismatches 49; Indels   0; Gaps   0;

Qy   12 TTTTTTTTTTTTTTTTAAGTAGAGAGGAGAGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 71
             ||||||||| | |||||||||||||| ||||||||||||||||||||||||||||
Db   72436 TTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACT 72495

Qy   72 CCTGACCTCGTGATCCGCGCCCCGGCCTTCCCAAAGTGCTGGGATTACAGGCGCCATGAGCC 131
             |||||||||||||||||||||| ||||| ||||||||| |||||||| ||||||| |||||
Db   72496 CCTGACCTCGTGATCCGCGCCCTCCAAAGTGTTGGGATTACAGGCGTGAGCC 72555

Qy   132 ACTGCGCCCAGCGCGGTCTTTTAAACATTCCCAGGACTGTACAGCCAACCCATACTCAC 191
             || ||||| | | |||| | | || ||| | || ||| | | | ||||||
Db   72556 ACCACGCCCGGGCCTTACATTTTTCATGATCGGGCAAGAAAAGTTGAACACTTGTAAC 72615

Qy   192 CTGACCATTTGGGAACTCCCCCCCCACGGCCCATA 223
             || |||| | ||| | ||| || ||| || |||
Db   72616 CTAACATGTGGGGATATTCCACCAATTGGTCAAA 72647

RESULT 7
BX000688/c
LOCUS   BX000688   102684 bp   DNA   linear   PRI 24-JUN-2003
DEFINITION Human DNA sequence from clone DAQB-36F16 on chromosome 6, complete
            sequence.
ACCESSION   BX000688
VERSION   BX000688.11 GI:32187946
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 102684)
  AUTHORS   Almeida,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (24-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Jun 24, 2003 this sequence version replaced gi:32168690.

------------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            DAQB-36F16 is from a DNA-arts QBL human bac library VECTOR:
            pBeloBAC11.
FEATURES   Location/Qualifiers
     source   1..102684
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAQB-36F16"
                     /clone_lib="DNA-arts-BAC.1-QBL.1"
ORIGIN

Query Match   15.4%; Score 133.6; DB 9; Length 102684;
Best Local Similarity   93.9%; Pred. No. 3.2e-29;
Matches 139; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTAAGTAGAGATGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
             |||||||||||||||||| |||||| ||||| ||||||||||||||||||||||||||||||
Db   84698 TTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 84639

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
             |||||||||||||||||||||||| ||||||||||| |||||||||||||||||||||||
Db   84638 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAATGCTGGGATTACAGGCG 84579

Qy   126 TGAGCCACTGCGCCCAGCGCGGTCTT 153
             ||||||||||||||||||||||| |||
Db   84578 TGAGGCCACCGGCGCGCCCAGCCCTTTTTTTT 84551

RESULT 8
AC006137
LOCUS   AC006137   129806 bp   DNA   linear   PRI 19-OCT-2001
DEFINITION Homo sapiens clone SCb-254N2 (UWGC:rg254N02) from 6p21, complete
            sequence.
ACCESSION   AC006137
VERSION   AC006137.3 GI:16271961
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 129806)
  AUTHORS   Janer,M., Guillaudeux,T., Vu,Q., Kutyavin,T., Harter,H. and
            Geraghty,D.B.

TITLE   Large scale sequence analysis of the human MHC class I region
JOURNAL   Unpublished
REMARK   Fred Hutchinson Cancer Research Center
          The Clinical Research Division
          1100 Fairview Ave. N., P.O. Box 19024
          Seattle, WA 98109-1024

REFERENCE 2 (bases 1 to 129806)
AUTHORS   Geraghty,D.E. and Olson,M.V.
TITLE   Direct Submission
JOURNAL   Submitted (05-DEC-1998) Human Genome Center, University of
          Washington, Box 352145, Seattle, WA 98195, USA
REMARK   University of Washington Human Genome Center
          Box 352145, Seattle, WA 98195
          Contact: Daniel E. Geraghty (geraghty@fhcrc.org)

REFERENCE 3 (bases 1 to 129806)
AUTHORS   Kaul,R.K. and Haugen,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (24-AUG-2000) Genome Center, University of Washington,
          Box 352145, Seattle, WA 98195, USA
REMARK   University of Washington Genome Center
          Box 352145 Seattle, WA 98195
          Contact: Rajinder Kaul (rkkaul@u.washington.edu)

REFERENCE 4 (bases 1 to 129806)
AUTHORS   Kaul,R.K. and Haugen,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (19-OCT-2001) Genome Center, University of Washington,
          Box 352145, Seattle, WA 98195, USA
REMARK   On Oct 19, 2001 this sequence version replaced gi:9987710.

COMMENT   Sequence Validation:
          This sequence has been validated by Multiple Complete Digest
          Mapping. Comparison of the experimentally derived map digest
          fragments with sequence-predicted fragments is given below.
          Small fragments below a variable cutoff (approximately 400-600bp)
          are not mapped and hence do not appear in the table. There are no
          significant remaining discrepancies between the experimental and
          predicted values. Uniquely ordered fragment groups are separated
          by dashed lines.

BglII Map	Seq	HindIII Map	Seq	NsiI Map	Seq
6695.83	6654.00	3764.58	3751.00	3143.81	3146.00
2829.33	2792.00	2619.67	2605.00	7632.25	7732.00
---	---	---	---	---	---
731.67	723.00	1182.00	1181.00	4157.75	4197.00
1451.30	1461.00	4416.83	4379.00	754.07	745.00
3011.22	2991.00	12731.50	12909.00	2612.60	2609.00
8308.95	8430.00	4045.43	4133.00	6018.56	6104.00
5483.63	5517.00	1708.46	1697.00	5504.56	5611.00
1411.52	1408.00	2824.85	2828.00		
3036.94	3067.00	3283.08	3301.00	3646.39	3681.00
				3796.27	3886.00

BglII Map	Seq	HindIII Map	Seq	NsiI Map	Seq
739.54	740.00	2519.25	2440.00	5115.00	5097.00
3935.08	3908.00	4420.00	4411.00	6536.33	6621.00
6360.62	6319.00	5049.00	5001.00	5353.67	5202.00
4507.12	4455.00	1033.12	1012.00	10074.25	9987.00
1288.62	1303.00	2099.75	2064.00	1747.83	1751.00
12213.75	12356.00	17212.25	18083.00	10342.00	10215.00
2793.33	2679.00	10608.42	10467.00	4666.08	4599.00
4379.83	4328.00	1932.92	1925.00	2849.83	2850.00
1286.75	1280.00	12731.50	12500.00	1828.83	1831.00
879.17	863.00	1584.50	1580.00	3687.17	3722.00
1816.33	1826.00	819.00	788.00	4181.75	4110.00
11414.25	11761.00	2619.67	2569.00	10257.92	10055.00
6029.33	5933.00	2619.67	2581.00	6413.17	6349.00
6029.33	6005.00	953.00	954.00	1371.00	1370.00
5273.83	5308.00	1702.83	1697.00	1132.00	1121.00
15738.17	15951.00	1244.08	1238.00	5439.00	5587.00

FEATURES   Location/Qualifiers
     source   1..129806
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21"
                     /clone="UWGC:rg254N02"
                     /clone_lib="Research Genetics"
     misc_feature   24682..24683
                     /note="IS186 and target-site duplication removed from
                     between bases 24682 and 26020"

ORIGIN

Query Match   15.4%; Score 133.6; DB 9; Length 129806;
  Best Local Similarity   93.9%;
  Matches 139; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              ||||||||||||||||||| ||||||||||||||| ||||||||||||||| |||||||||||||
Db   40206 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTGCCAGGATGGTCT 40265

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
              ||| |||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   40266 CGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAATGCTGGGATTACAGGCG 40325

Qy   126 TGAGCCACTGCGCCCAGCCGGGTCTTTTTT 153
              |||||||| ||||| ||||||| ||||||
Db   40326 TGAGCCACCGCGCCGCCCAGCCTTTTTTTTTT 40353

RESULT 9
LOCUS   HS271M21   134292 bp   DNA   linear   PRI 14-SEP-2001
DEFINITION Human DNA sequence from clone RP1-271M21 on chromosome
            6p21.31-22.2. Contains a MAS1 oncogene pseudogene, the OR2H2 gene
            for olfactory receptor 2H2, olfactory receptor 2I1 and 2H5
            pseudogenes OR2I1P and OR2H5P, the gene for ubiquitin, an RPL13A
            (60S Ribosomal Protein L13A) pseudogene, the GABBR1 gene for
            gamma-aminobutyric acid (GABA) B receptor 1 and an SMT3H1 or SMT3H2
            (SMT3 (suppressor of mif two 3, yeast) homolog) pseudogene.

```
ACCESSION   AL031983
VERSION     AL031983.2  GI:12329558
KEYWORDS    HTG; CpG island; diubiquitin; GABBR1; olfactory receptor; OR2H2;
            OR2H5P; OR2I1P; RPL13A; SMT3H1; SMT3H2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 134292)
  AUTHORS   Younger,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 21, 2001 this sequence version replaced gi:4160199.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-271M21 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            This sequence was finished as follows unless otherwise noted; all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. IMPORTANT: This
            sequence is not the entire insert of clone RP1-271M21 It may be
            shorter because we sequence overlapping sections only once, except
            for a 100 base overlap.
            The true left end of clone RP1-271M21 is at 134292 in this
            sequence. The true right end of clone RP1-271M21 is at 1 in this
            sequence.
FEATURES             Location/Qualifiers
     source          1..134292
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p21.31-22.2"
                     /clone="RP1-271M21"
                     /clone_lib="RPCI-1"
     repeat_region   83..213
                     /note="MER94 repeat: matches 1..134 of consensus"
     repeat_region   518..879
                     /note="MLT1A1 repeat: matches 1..365 of consensus"
     repeat_region   5038..5466
                     /note="MLT1C repeat: matches 17..466 of consensus"
     misc_feature    5109..5561
                     /note="match: GSS: Em:AQ435929"
     repeat_region   5815..5853
                     /note="MIR repeat: matches 108..146 of consensus"
     repeat_region   7324..7786
                     /note="HERV16 repeat: matches 596..1091 of consensus"
     repeat_region   7857..8024
                     /note="L1M4 repeat: matches 3415..3586 of consensus"
     repeat_region   8191..8565
                     /note="THE1C repeat: matches 1..371 of consensus"
     repeat_region   8604..8910
                     /note="AluSx repeat: matches 2..306 of consensus"
     repeat_region   9002..9354
                     /note="MLT1B repeat: matches 1..390 of consensus"
     repeat_region   9717..9795
                     /note="MIR repeat: matches 7..167 of consensus"
     repeat_region   10027..10047
                     /note="L1MB6 repeat: matches 20..98 of consensus"
     repeat_region   10048..10493
                     /note="MLT1C repeat: matches 5767..5786 of consensus"
     repeat_region   10494..10848
                     /note="L1MB6 repeat: matches 5786..6145 of consensus"
     repeat_region   11216..11356
                     /note="L2 repeat: matches 2554..2703 of consensus"
     repeat_region   11386..11537
                     /note="MIR repeat: matches 11786..11937 of consensus"
     repeat_region   11610..11786
                     /note="L1MC2 repeat: matches 4759..4931 of consensus"
     repeat_region   11787..12095
                     /note="L1MC2 repeat: matches 4931..4931 of consensus"
     repeat_region   12096..13044
                     /note="AluSg1 repeat: matches 1..308 of consensus"
     repeat_region   13099..13476
                     /note="L1MC2 repeat: matches 5728..6118 of consensus"
     repeat_region   13916..13991
                     /note="L1MB1 repeat: matches 69..147 of consensus"
     repeat_region   14700..14834
                     /note="MIR repeat: matches 89..226 of consensus"
     repeat_region   16392..16688
                     /note="MER9 repeat: matches 215..511 of consensus"
     repeat_region   16994..17015
                     /note="11 copies 2 mer aa 100% conserved"
     misc_feature    17498..17907
                     /note="match: GSS: Em:AQ883063"
     misc_feature    17524..17673
                     /note="match: GSS: Em:B724414"
     repeat_region   17750..18033
                     /note="MER50 repeat: matches 1..307 of consensus"
     misc_feature    18059..18968
                     /note="match: GSS: Em:AQ743745"
     repeat_region   18234..19882
                     /note="MER4-internal repeat: matches 2701..4800 of
                     consensus"
     repeat_region   19883..20181
                     /note="LTR36 repeat: matches 1..612 of consensus"
     repeat_region   20182..20687
                     /note="AluJb repeat: matches 1..292 of consensus"
     repeat_region   20675..20796
                     /note="MER4-internal repeat: matches 2198..2701 of
                     consensus"
     repeat_region   20777..20853
                     /note="MER4-internal repeat: matches 2341..2464 of
                     consensus"
     misc_feature    20854..21376
                     /note="match: GSS: Em:B712414"
     repeat_region   21377..22220
                     /note="MER4-internal repeat: matches 121..1037 of
                     consensus"
     repeat_region   22222..22934
                     /note="MER50 repeat: matches 1..711 of consensus"
     repeat_region   22959..23269
                     /note="AluJb repeat: matches 1..307 of consensus"
     repeat_region   23272..23353
                     /note="MER4-internal repeat: matches 1..84 of consensus"
     repeat_region   23357..23776
                     /note="MER4C repeat: matches 1..465 of consensus"
     misc_feature    26589..27225
                     /note="match: GSS: Em:AQ195255"
     repeat_region   28927..29156
                     /note="STS: Em:G61733"
     repeat_region   30633..30690
                     /note="SN5 repeat: matches 1..235 of consensus"
```

repeat_region   31019..31135
                 /note="29 copies 2 mer gt 74% conserved"
repeat_region   31856..32261
                 /note="AluJb repeat: matches 1..311 of consensus"
repeat_region   32517..32893
                 /note="L1M4 repeat: matches 4650..5281 of consensus"
repeat_region   32899..33125
                 /note="L1M4 repeat: matches 3785..4144 of consensus"
misc_feature   33093..33354
                 /note="L1MEc repeat: matches 1787..2016 of consensus"
                 /note="match: GSS: Em:AQ620121"
gene   complement(34169..35034)
                 /gene="dJ271M21.8"
                 /pseudo
repeat_region   33468..33574
                 /note="MIR repeat: matches 144..252 of consensus"
gene   33679..34067
                 /gene="dJ271M21.8"
                 /pseudo
repeat_region   33207..33337
                 /note="L1M4 repeat: matches 4708..4838 of consensus"
repeat_region   33467..33574
                 /note="MSTA repeat: matches 1..426 of consensus"
CDS   complement(34169..35034)
                 /gene="dJ271M21.8"
                 /pseudo
                 /codon_start=1
                 /evidence=not_experimental
                 /note="dJ271M21.8 (MAS1 oncogene pseudogene)
                 match: proteins: Sw:P30554 Sw:P12526 Sw:P04201 Sw:P35410"
repeat_region   35631..35754
                 /note="L2 repeat: matches 2367..2511 of consensus"
repeat_region   36169..36279
                 /note="MER81 repeat: matches 3..113 of consensus"
repeat_region   36465..36829
                 /note="MIR1B repeat: matches 1..374 of consensus"
repeat_region   37958..38299
                 /note="L1MB5 repeat: matches 5817..6165 of consensus"
repeat_region   39291..39590
                 /note="AluSp repeat: matches 1..301 of consensus"
repeat_region   40192..41538
                 /note="L1M4 repeat: matches 2220..3639 of consensus"
repeat_region   41538..41700
                 /note="L1M4c repeat: matches 1140..1302 of consensus"
repeat_region   41701..42011
                 /note="AluYa5 repeat: matches 1..311 of consensus"
repeat_region   42012..42297
                 /note="L1M4c repeat: matches 1301..2216 of consensus"
repeat_region   43281..43540

RESULT 10
LOCUS   AL662826   145431 bp   DNA   linear   PRI 22-MAY-2002
DEFINITION Human DNA sequence from clone XXbac-101120 on chromosome 6,
            complete sequence.
ACCESSION   AL662826
VERSION   AL662826.11 GI:20870241
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)

Query Match   15.4%;   Score 133.6;   DB 9;   Length 134292;
Best Local Similarity   93.9%;   Pred. No. 3.3e-29;
Matches 139;   Conservative 0;   Mismatches 9;   Indels 0;   Gaps 0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTAAGTAGTAGAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113399 TTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 113340

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113339 CGATCTCCTGACCTCGTGATCCCGCCCGCCTCCCAAAATGCTGGGATTACAGGCG 113280

Qy   126 TGAGCCACTGCGCCCAGCCGGGTCTTTTT 153
              ||||||||||||||||||||||||||||||
Db   111339 TGAGCCACCGCGCCCAGCCTTTTTTTTTTT 111252

ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 145431)
  AUTHORS   Brown,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (22-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On May 16, 2002 this sequence version replaced gi:20793440.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            XXbac-101120 is from a CHORI-502 human bac - COX cell line library
            VECTOR: pTARBAC2.1.
FEATURES   Location/Qualifiers
     source   1..145431
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="XXbac-101120"
                     /clone_lib="CHORI-502"

RESULT 11
LOCUS   AP006245   155586 bp   DNA   linear   PRI 04-MAR-2003
DEFINITION Homo sapiens genomic DNA, chromosome 8, clone:RP11-139F9, complete
            sequence.
ACCESSION   AP006245
VERSION   AP006245.1 GI:28812033
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Query Match   15.4%;   Score 133.6;   DB 9;   Length 145431;
Best Local Similarity   93.9%;   Pred. No. 3.3e-29;
Matches 139;   Conservative 0;   Mismatches 9;   Indels 0;   Gaps 0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTAAGTAGTAGAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91119 TTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 91060

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   91059 CGATCTCCTGACCTCGTGATCCCGCCCGCCTCCCAAAATGCTGGGATTACAGGCG 91000

Qy   126 TGAGCCACTGCGCCCAGCCGGGTCTTTTT 153
              ||||||||||||||||||||||||||||||
Db   90999 TGAGCCACCGCGCCCAGCCTTTTTTTTTTT 90972

REFERENCE   1
AUTHORS   Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
             Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE   Homo sapiens genomic DNA
JOURNAL   Published Only in Database (2003)
REFERENCE   2 (bases 1 to 155586)
AUTHORS   Hattori,M. Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
             Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (27-FEB-2003) Masahira Hattori, The Institute of Physical
             and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
             1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
             Tel:81-45-503-9111, Fax:81-45-503-9170)
FEATURES   Location/Qualifiers
     source   1..155586
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /clone="RP11-139F9"
ORIGIN

Query Match   15.4%; Score 133.6; DB 9; Length 155586;
Best Local Similarity   76.9%; Pred. No. 3.3e-29;
Matches 163; Conservative   0; Mismatches 49; Indels   0; Gaps   0;

Qy   12 TTTTTTTTTTTTTTTAAGTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 71
          ||||||||| ||| ||||||| ||||||||||||||||||||||||||||||||||||
Db 82858 TTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCCAATCT 82799

Qy   72 CCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCC 131
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 82798 CCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCCATGAGCC 82739

Qy   132 ACTGCGCCCAGCCGCGGTCTTTTTAAACATTCCCAGGACTGTTACAGCCAACCCATACTCAC 191
          |||||||||||||||| |||||||| |||||||||||||||||||||||||||| ||||
Db 82738 ACCACGCCCGCCTACATTTTCATGATCCATAGGGCAAGAAAAGTTGAACACTTGTAAC 82679

Qy   192 CTGACATTTGGGAAGTCCCCCCCCACGGCCCATA 223
          ||||||||||||||| ||||||||||||| |||
Db 82678 CTAACATGTGGGATATTCCACCATTGGTCAAA 82647

RESULT 12
LOCUS   AL645936/c
ACCESSION   AL645936
DEFINITION   Human DNA sequence from clone Xxbac-126D10 on chromosome 6,
             complete sequence.
ACCESSION   AL645936
VERSION   AL645936.5 GI:19572781
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Leongamornlert,D.
TITLE   Direct Submission
JOURNAL   Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
             Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
             humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Mar 21, 2002 this sequence version replaced gi:19572047.
             During sequence assembly data is compared from overlapping clones.
             Where differences are found these are annotated as variations
             together with a note of the overlapping clone name. Note that the
             variation annotation may not be found in the sequence submission
             corresponding to the overlapping clone, as we submit sequences with
             only a small overlap as described above.
             This sequence was finished as follows unless otherwise noted: all
             regions were either double-stranded or sequenced with an alternate
             chemistry or covered by high quality data (i.e., phred quality >=
             30); an attempt was made to resolve all sequencing problems, such

LOCUS   AL645936   155874 bp   DNA   linear   PRI 24-APR-2002

as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep WORMPEP
from a CHORI-501 human bac - PGP cell line library VECTOR:
pTARBAC2.1
This sequence was generated from part of bacterial clone contigs
constructed by the MHC Haplotype Consortium and collaborators.
Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC.

FEATURES   Location/Qualifiers
     source   1..155874
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="Xxbac-126D10"
                     /clone_lib="CHORI-501"
ORIGIN

Query Match   15.4%; Score 133.6; DB 9; Length 155874;
Best Local Similarity   93.9%; Pred. No. 3.3e-29;
Matches 139; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTAAGTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 67990 TTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCT 67931

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 67930 CGATCTCCTGACCTCGTGATCCCCCGCCTCGGCCTCCCAAAATGCTGGGATTACAGGCG 67871

Qy   126 TGAGCCACTGCGCCCCAGCGCGGGTTCTTTTT 153
          |||||||||| ||||||| ||||| | |||||
Db 67870 TGAGCCACCGCGCGCCCAGCCCTTTTTTTT 67843

RESULT 13
LOCUS   AC138080/c
ACCESSION   AC138080
DEFINITION   Homo sapiens chromosome 8, clone RP11-1147M13, complete sequence.
ACCESSION   AC138080
VERSION   AC138080.4 GI:28565756
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 164872)
AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
TITLE   Homo sapiens chromosome 8, clone RP11-1147M13
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 164872)
AUTHORS   Birren,B., Nusbaum,C., Ali,A., Allen,N., Anderson,S.,
             Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
             Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
             Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
             Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
             Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
             Hagos,B., Horton,L., Hulme,W., Iiley,I., Johnson,R., Jones,C.,
             Kamat,A., Karalas,A., Kells,C., Landers,T., Levine,R.,
             Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
             Mathews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
             Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
             Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
             Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
             Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R.,
             Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
             Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,

```
TITLE       Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
              Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
JOURNAL     Direct Submission
            Submitted (12-DEC-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 164872)
AUTHORS     Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
              Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
              Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
              Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
              Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
              Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
              Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
              Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
              Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
              Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
              Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
              Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
              Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
              Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
              Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
              Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
              Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (15-JAN-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 164872)
AUTHORS     Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
              Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
              Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
              Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
              Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
              Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
              Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
              Hall,J., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
              Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
              Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
              Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
              Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
              Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
              O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
              Rachupka,A., Ramaswamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
              Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
              Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
              Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
              Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
              Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (26-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 26, 2003 this sequence version replaced gi:27753762.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            --------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            --------- Project Information
            Center project name: L28708
            Center clone name: 1147_M_13
            ---------
FEATURES             Location/Qualifiers
     source          1..164872
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="RP11-1147M13"
                     /clone_lib="RPCI-11 Human Male BAC"
```

```
     repeat_region   862..1167
                     /rpt_family="AluJo"
     repeat_region   1509..1629
                     /rpt_family="MER5A"
     repeat_region   1650..1880
                     complement(2025..2158)
                     /rpt_family="MIR"
     repeat_region   2430..2627
                     complement(3246..3459)
                     /rpt_family="MLT1J"
     repeat_region   3462..3489
                     /rpt_family="(A)n"
     repeat_region   3779..4068
                     /rpt_family="AluSc"
     repeat_region   4113..4147
                     complement(4269..4521)
                     /rpt_family="(CA)n"
     repeat_region   4526..4549
                     /rpt_family="LTR50"
     repeat_region   4558..4866
                     /rpt_family="(CAAAA)n"
     repeat_region   5713..6040
                     complement(5001..5111)
                     /rpt_family="AluB"
     repeat_region   6534..6688
                     /rpt_family="LTR50"
     repeat_region   7856..7900
                     complement(7781..7854)
                     /rpt_family="MLT2B1"
     repeat_region   7911..8055
                     /rpt_family="(CAGAGA)n"
     repeat_region   8022..8069
                     /rpt_family="(TA)n"
     repeat_region   8557..8798
                     complement(8072..8454)
                     /rpt_family="MLT2B1"
     repeat_region   9121..9258
                     /rpt_family="MIR"
     repeat_region   10103..10321
                     complement(10103..10321)
                     /rpt_family="FRAM"
     repeat_region   10406..10776
                     /rpt_family="MIR"
     repeat_region   10812..11181
                     /rpt_family="MLT1A"
     repeat_region   11133..11463
                     /rpt_family="LTR37B"
     repeat_region   11969..12278
                     /rpt_family="L2"
     repeat_region   13168..13252
                     complement(13168..13252)
                     /rpt_family="AluY"
     repeat_region   13306..13382
                     /rpt_family="MIR"
     repeat_region   13642..13933
                     complement(14209..14299)
                     /rpt_family="MIR"
     repeat_region   15413..15567
                     /rpt_family="AluSx"
     repeat_region   15568..15845
                     complement(15877..16089)
                     /rpt_family="(TTTC)n"
     repeat_region   16405..16623
                     /rpt_family="AluSx"
     repeat_region   16794..16843
                     /rpt_family="L1ME"
     repeat_region   17367..17628
                     /rpt_family="(TA)n"
```

```
RESULT 14
ACC97329
LOCUS       AC097329              199706 bp    DNA     linear   HTG 20-OCT-2001
DEFINITION  Pan troglodytes clone RP43-45G20, WORKING DRAFT SEQUENCE, 10
            unordered pieces.
ACCESSION   AC097329
VERSION     AC097329.1  GI:16118152
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 199706)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.P.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Deaderick,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,B., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,X., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199706)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     ---------------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ---------------------- Project Information
            Center project name: ZUAO
            Center clone name: RP43-45G20
            ---------------------- Summary Statistics
            Sequencing vector: Plasmid; M77789
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 195498 bases at least Q40
            Consensus quality: 201324 bases at least Q30
            Consensus quality: 206344 bases at least Q20
            Estimated insert size: 207444; sum-of-contigs estimation
            Quality coverage: 0x in Q20 bases; agarose-fp estimation
            Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
            ------------------------------------------------------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            *      1           47628: contig of 47628 bp in length
            *  47629           47728: gap of unknown length
            *  47729          102375: contig of 54647 bp in length
            * 102376          102475: gap of unknown length
            * 102476          115011: contig of 12536 bp in length
            * 115012          115111: gap of unknown length
            * 115112          139605: contig of 24494 bp in length
            * 139606          139705: gap of unknown length
            * 139706          160858: contig of 21153 bp in length
            * 160859          160958: gap of unknown length
            * 160959          169439: contig of 8481 bp in length
```

```
                               Query Match
Best Local Similarity   15.4%;    Score 133.6; DB 9;   Length 164872;
Matches 163; Conservative   76.9%;    Pred. No. 3.3e-29;
               0; Mismatches    49;   Indels    0;   Gaps    0;

Qy      12  TTTTTTTTTTTTTTAAGTAGAGANATGGGGTTTCACCGTGTTAGCCAAGGANTGGTCTTCGAACT    71
            ||||||||||||||| || ||||||||||||||||||||||| || |||||||||||||| |||
Db   12101  TTTTTTTTGTANTTTTAGTAGAGACGGGGTTTCACCCGTGTTAGCCAGGATGGTCTCCATCT  12042

Qy      72  CCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC   131
            ||| |||||||||||||||||||||||||||||||||| |||||||||||||||| ||||
Db   12041  CCTGACCTCGTGATCCGCCCGCCTCAGCCTCCCAAAGTGTTGGGATTACAGGCGTGAGCC  11982

Qy     132  ACTGCGCCCAGCCGGTCTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCAC   191
            || ||| |||   |||||||||||||||||||||||||||||||||||||||||||||
Db   11981  ACCACGCCCGGCCTACACATTTTCATCGANGCATAGGGCCANGAAAAGTTGAACACTTGTAAC  11922

Qy     192  CTGACATTTGGGAACTCCCCCCCCCACGGCCATA   223
            ||  |||  ||| ||||||||||||||||||
Db   11921  CTAACATGTGGGGATATCTCACCATTGGTCAAA  11890
```

```
repeat_region  17680..18076        /rpt_family="LlMC4a"
repeat_region  18317..18351        /rpt_family="LlMC4a"
repeat_region  18402..18703        /rpt_family="(CAAA)n"
repeat_region  18704..18859        /rpt_family="AluSx"
repeat_region  19224..19419        /rpt_family="AluSg/x"
repeat_region  19558..19593        /rpt_family="HAL1"
repeat_region  19627..19696        /rpt_family="HAL1"
repeat_region  19697..20000        /rpt_family="(CA)n"
repeat_region  20001..20267        /rpt_family="HAL1"
repeat_region  20487..20666        /rpt_family="AluY"
               complement(20714..21381)  /rpt_family="HAL1"
repeat_region  complement(21663..21730)  /rpt_family="L1PA3"
```

RESULT 15

```
LOCUS       AL355484              207362 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 1 clone RP11-370E21, 9 unordered pieces.
ACCESSION   AL355484
VERSION     AL355484.10  GI:10039826
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mclay.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 9, 2000 this sequence version replaced gi:9926590.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA370E21
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator ET-amersham; 31% of reads Chemistry:
            Dye-terminator Big Dye; 68% of reads
            Consensus quality: 202124 bases at least Q40
            Consensus quality: 204004 bases at least Q30
            Consensus quality: 205180 bases at least Q20
            Insert size: 206562; sum-of-contigs
            Insert size: 201559; 10.0% error; agarose-fp
            Quality coverage: 5.18x in Q20 bases; sum-of-contigs Quality
            coverage: 5.38x in Q20 bases; agarose-fp
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
```

```
FEATURES             Location/Qualifiers
     source          1..199706
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-45G20"
ORIGIN
Query Match         15.4%;   Score 133.6;  DB 2;  Length 199706;
Best Local Similarity   86.0%;   Pred. No.3.3e-29;
Matches 148;   Conservative    0;   Mismatches  24;   Indels   0;   Gaps    0;

Qy        2 CGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAAGAGGATGGGGGTTTCACCGTGTTAGCCAGGATG 61
            |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   102755 CTAATTTTTTTTTTTTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATG 102814

Qy       62 GTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACA 121
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   102815 GTCTCGAACTCCTGAACTCATCACCGCGCCTCAAAGTGCTGGGATTACA 102874

Qy      122 GGCATGAGCCACTCGCCCGCCCAGCCGGTCTTTTTAAACATTCCCCAGGACTGTA 173
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   102875 GGCATGAGCCACTCGCCCGCCCGGCCGAGTAATGTTAATTTTTCTCAAAACTGAA 102926
```

FEATURES
```
     source          1..207362
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-370E21"
                     /clone_lib="RPCI-11.2"
     misc_feature    1..7932
                     /note="assembly fragment:00060"
     misc_feature    8033..20003
                     /note="assembly fragment:chain:1"
     misc_feature    20104..23088
                     /note="assembly fragment:01488"
     misc_feature    23189..49409
                     /note="assembly fragment:chain:1"
     misc_feature    49510..54141
                     /note="assembly fragment:01485"
     misc_feature    54242..104573
                     /note="assembly fragment:chain:1"
     misc_feature    54442..104573
                     /note="assembly fragment:02543"
     misc_feature    104674..127553
                     /note="assembly fragment:00004"
     misc_feature    104674..127553
                     /note="assembly fragment:02462"
     misc_feature    127654..131260
                     /note="assembly fragment:chain:2"
     misc_feature    131361..207362
                     /note="assembly fragment:02394"
     misc_feature    131361..207362
                     /note="assembly fragment:chain:2"
                     /note="assembly fragment:00303"
                     clone_end:SP6
                     vector_side:right"
```

ORIGIN
Query Match   15.4%;   Score 133.6; DB 2; Length 207362;
Best Local Similarity   76.9%;   Pred. No.3.3e-29;
Matches 163;   Conservative   0;   Mismatches 49;   Indels   0;   Gaps   0;

```
Qy       12 TTTTTTTTTTTTTTAAGTAAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACT 71
            |||||||||||| | ||||||||| |||||||||||||||||||||||||||||||||
Db    35791 TTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCCATCT 35850

Qy       72 CCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC 131
            ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    35851 CCTGACCTCGTGATCCGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC 35910
```

The following FEATURES and source block appears at top of the first column:

```
FEATURES             Location/Qualifiers
     source          1..199706
          *   169440   169539: gap of unknown length
          *   169540   182344: contig of 12805 bp in length
          *   182345   182444: gap of unknown length
          *   182445   190234: contig of 7791 bp in length
          *   190235   190335: gap of unknown length
          *   190336   196663: contig of 6328 bp in length
          *   196664   196763: gap of unknown length
          *   196764   199705: contig of 2943 bp in length.
          *   199706: gap of unknown length
```

```
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
     source          1..207362
          *       1    7932: contig of 7932 bp in length
          *    7933    8032: gap of 100 bp
          *    8033   20003: contig of 11971 bp in length
          *   20004   20103: gap of 100 bp
          *   20104   23088: contig of 2985 bp in length
          *   23089   23188: gap of 100 bp
          *   23189   49409: contig of 26221 bp in length
          *   49410   49509: gap of 100 bp
          *   49510   54141: contig of 4632 bp in length
          *   54142   54241: gap of 100 bp
          *   54242  104573: contig of 50332 bp in length
          *  104574  104673: gap of 100 bp
          *  104674  127553: contig of 22880 bp in length
          *  127554  127653: gap of 100 bp
          *  127654  131260: contig of 3607 bp in length
          *  131261  131360: gap of 100 bp
          *  131361  207362: contig of 76002 bp in length.
```

```
Qy     132  ACTGCGCCCAGCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCATACTCAC 191
            | || ||||| || | |  |||||||| |  |  |  | | |  ||  | | || |
Db   35911  ACCACGCCCCGGCCCTACATTTTTCATGATGCATAGGGGCAAGAAAAGTTGAACACTTGTAAC 35970

Qy     192  CTGACATTTGGGAACTCCCCCCCACGGGCCATA 223
            || |||||||| ||||||| | |   || | |
Db   35971  CTAACATGTGGGATATCTCACCATTGGTCAAA 36002
```

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 15:50:42 ; Search time 2500 Seconds
(without alignments)
15031.364 Million cell updates/sec

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Title: US-10-621-363-12
Perfect score: 867
Sequence: 1 tcggtttttttttttttt...........accaatctcgtgtcgaattc 867

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 2167516995 residues

Word size :   0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	483	55.7	124874	9	HS888M10	AL031296 Human DNA
c 2	269	31.0	2069	9	AK095438	AK095438 Homo sapi
c 3	98	11.3	145831	6	AX330137	AX330137 Sequence
c 4	98	11.3	145831	6	AX334634	AX334634 Sequence
c 5	98	11.3	145831	6	AX336416	AX336416 Sequence
c 6	98	11.3	145831	6	AX337041	AX337041 Sequence
c 7	98	11.3	145831	6	HUAF001548	HUAF001548 Human Chr
c 8	98	11.3	202014	9	BS000050	BS000050 Pan trogl
c 9	89	10.3	75555	9	AC021026	AC021026 Homo sapi
c 10	89	10.3	88872	9	AC045537	AC045537 Homo sapi
c 11	89	10.3	149397	9	AC004840	AC004840 Homo sapi
c 12	89	10.3	152005	9	AC020890	AC020890 Homo sapi
c 13	89	10.3	184927	2	AL683875	AL683875 Homo sapi
c 14	89	10.3	189825	9	AL683807	AL683807 Human DNA
c 15	86	9.9	57447	9	AL355146	AL355146 Human DNA
c 16	86	9.9	60442	9	AP001342	AP001342 Homo sapi
c 17	86	9.9	158146	9	AC005076	AC005076 Homo sapi.
c 18	86	9.9	168513	2	AC011982	AC011982 Homo sapi.
c 19	86	9.9	173117	2	AL157885	AL157885 Homo sapi
c 20	86	9.9	340000	9	AP001692	AP001692 Homo sapi.
c 21	82	9.5	407	9	AY053360	AY053360 Homo sapi.
c 22	82	9.5	43454	9	AC002056	AC002056 Homo sapi.
c 23	79	9.1	21579	9	AC138417	AC138417 Homo sapi.
c 24	79	9.1	145540	9	AP001052	AP001052 Homo sapi.
c 25	79	9.1	167672	2	AC132806	AC132806 Homo sapi.
c 26	79	9.1	177719	9	AC135048	AC135048 Homo sapi.
c 27	79	9.1	340000	2	AP001752	AP001752 Homo sapi.
c 28	78	9.0	98935	9	AC004935	AC004935 Homo sapi.
c 29	78	9.0	154752	9	AC108112	AC108112 Homo sapi.
c 30	78	9.0	163100	2	AC145421	AC145421 Homo sapi.
c 31	78	9.0	173786	2	AC024472	AC024472 Homo sapi.
c 32	78	9.0	185758	2	AC026552	AC026552 Homo sapi.
c 33	77	8.9	80552	9	AL353377	AL353377 Human DNA
c 34	77	8.9	151136	9	HSA189K21	AL121893 Human DNA
c 35	76	8.8	60375	9	AL355679	AL355679 Human DNA
c 36	76	8.8	87434	9	HSA297357	AJ297357 Homo sapi.
c 37	76	8.8	159887	2	AC067823	AC067823 Homo sapi.
c 38	76	8.8	165261	2	AC012659	AC012659 Homo sapi.
c 39	76	8.8	176741	9	AC098476	AC098476 Homo sapi.
c 40	76	8.8	189993	9	AP002762	AP002762 Homo sapi.
c 41	75	8.7	54649	9	AL135935	AL135935 Human DNA
c 42	75	8.7	94414	9	AC051663	AC051663 Homo sapi.
c 43	75	8.7	170787	9	AC121247	AC121247 Homo sapi.
c 44	75	8.7	199103	2	AC080076	AC080076 Homo sapi.
c 45	75	8.7	203560	2	AC026298	AC026298 Homo sapi.

ALIGNMENTS

RESULT 1
LOCUS   HS888M10   124874 bp   DNA   linear   PRI   04-MAR-2003
HS888M10/c
DEFINITION   Human DNA sequence from clone RP5-888M10 on chromosome
     1p36.11-36.31, complete sequence.
ACCESSION   AL031296
VERSION   AL031296.1   GI:4106564
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
   ORGANISM   Homo sapiens
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 124874)
AUTHORS   Cobley,V.
TITLE   Direct Submission

JOURNAL   Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT   On Jan 5, 1999 this sequence version replaced gi:3881974.
--------------------------- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP. Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP5-888M10 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES   Location/Qualifiers
    source   1..124874
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="RZPD:RPCIP704M10888"
              /db_xref="taxon:9606"
              /chromosome="1"
              /map="p36.11-36.31"
              /clone="RP5-888M10"
              /clone_lib="RPCI-5"

ORIGIN

Query Match   55.7%; Score 483; DB 9; Length 124874;
Best Local Similarity 99.5%; Pred. No. 4.1e-273;
Matches 823; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```
Qy        6 TTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTCACGTGTTAGCCAGGATGGTCT 65
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87285 TTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACGTGTTAGCCAGGATGGTCT 87226

Qy       66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87225 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 87166

Qy      126 TGAGCCACTGCGCCCAGCCGGTCTTTTTAAACATTCCCCCAGGACTTGTACNGCCAACCCAT 185
            |||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db    87165 TGAGCCACTGCGCCCAGCCGGTCTTTTTAAACATTCCCCCAGGACTTGTACAGCCAACCCAT 87106

Qy      186 ACTCACCTGACATTTGGGAACTCCCCCCCCACCGGCCCATACTAACTGATCTGCAGAGGTAAGACC 245
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87105 ACTCACCTGACATTTGGGAACTCCCCCCCCACCGGCCCATACTAACTGATCTGCAGAGGTAAGACC 87046

Qy      246 AAGAGCAAGAAATGGGGGGATTCACATCTAAGGTCTGGTGATGGGCTGATGAAGGAAGAAGAA 305
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87045 AAGAGCAAGAAATGGGGGGATTCACATCTAAGGTCTGGTGATGGGCTGATGAAGGAAGAAGAA 86986
```


LOCUS   AK095438/c   2069 bp   mRNA   linear   PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ38119 fis, clone D3OST2004637.
ACCESSION   AK095438
VERSION   AK095438.1 GI:21754694
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2069)
  AUTHORS   Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
          Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
          Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
          Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
          Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
          Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
          Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
          Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
          Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE   NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 2069)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
          Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
          (E-mail:genomics@nhri.co.jp, Tel:81-438-52-3995, Fax:81-438-52-3986)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
          Economy, Trade and Industry of Japan; cDNA full insert sequencing:
          Research Association for Biotechnology (RAB); cDNA library
          construction: Helix Research Institute (HRI) (supported by Japan
          Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
          HRI, and Biotechnology Center, National Institute of Technology and
          Evaluation; clone selection for full insert sequencing: HRI and

```
Qy      306 TCAGCGGACAAAAAGCCTCTTAGGTCTTTCTTACACAAACACCTTCTCTGGCCACCTGCTTT 365
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86985 TCAGCGGACAAAAAGCCTCTTAGGTCTTTCTTACACAAACACCTTCTCTGGCCACCTGCTTT 86926

Qy      366 GAAAGGGGCAGAAGTATAGTGGGCGA-GCTGCCCACCTGCTACAGTGAAGGGATCTGGAG 424
            |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    86925 GAAAGGGGCAGAAGTATAGTGGGCGAGGCTGCCCACCTGCTACAGTGAAGGGATCTGGAG 86866

Qy      425 AAATACTCACACTTTGAGGTGCTCCGCCCTCTCATCAGCCCAGCTCTACTTAGGCCAATG 484
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86865 AAATACTCACACTTTGAGGTGCTCCGCCCTCTCATCAGCCCAGCTCTACTTAGGCCAATG 86806

Qy      485 ACCCCACCGGGAG-CTTACACAGTYCAAACAACAGCCCAAATGCATTCAGTGAGCAGGGGAG 543
            ||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||
Db    86805 ACCCCACCGGGAGGCTTACACAGTACAAACAACAGCCCAAATGCATTCAGTGAGCAGGGGAG 86746

Qy      544 GCCAAAGGACTCCGGAGGAGAGGAGGCCCAATAAGGCTGGTGCTATTTCCGGATCCATAGAG 603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86745 GCCAAAGGACTCCGGAGGAGAGGAGGCCCAATAAGGCTGGTGCTATTTCCGGATCCATAGAG 86686

Qy      604 AGAGCAGAGGTGGGGCAGGCCCTTTGATTAAATGTATCATTCTTGAATGCAAAGTTCAAAA 663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86685 AGAGCAGAGGTGGGGCAGGCCCTTTGATTAAATGTATCATTCTTGAATGCAAAGTTCAAAA 86626

Qy      664 TCCGGGTATGCCCGGGTGAGAATGAGCAGGACTAACACCTGGGGTGTCATGGCAAGCCTCCA 723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86625 TCCGGGTATGCCCGGGTGAGAATGAGCAGGACTAACACCTGGGGTGTCATGGCAAGCCTCCA 86566

Qy      724 GGGCCGACTGGCCCAGAGACAGGATCCGCAAGAGAGCCTCTCCAGCCAGCCTCTGGTGCCCAACC 783
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86565 GGGCCGACTGGCCCAGAGACAGGATCCGCAAGAGAGCCTCTCCAGCCAGCCTCTGGTGCCCAACC 86506

Qy      784 ACTCGGGATTTGAACCCCGGCTCCTCAAGGTCACGTGTGTGTAGCCCTTGA 830
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86505 ACTCGGGATTTGAACCCCGGCTCCTCAAGGTCACGTGTGTGTAGCCCTTGA 86459
```

RAB; annotation: HRI and RAB.

FEATURES   Location/Qualifiers
     source   1..2069
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="D3OST2004637"
                     /cell_type="CD34+ Cells"
                     /clone_lib="D3OST2"
                     /note="Cloning vector: pME18SFL3-mRNA from CD34+ cells
                     after 3-days ODF induction.-primary culture, CD34+ Cells"
ORIGIN

Query Match   31.0%;   Score 269; DB 9; Length 2069;
Best Local Similarity   100.0%; Pred. No. 6.9e-147;
Matches 269;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   6 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGAGGAGGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   269 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGAGGAGGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 210

Qy   66 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   209 CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 150

Qy   126 TGAGCCACTGCGCCCAGCCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAAACCAT 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   149 TGAGCCACTGCGCCCAGCCCGGTCTTTTTAAACATTCCCCAGGACTGTACAGCCAAACCAT 90

Qy   186 ACTCACCTGACATTTGGGAACTCCCCCCCCACAGGGCCCCATAACTGATCTCAGAGGGTAAGACC 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   89 ACTCACCTGACATTTGGGAACTCCCCCCCCACAGGGCCCCATAACTGATCTCAGAGGGTAAGACC 30

Qy   246 AAGAGCAAGAATGGGGGATTCACATCTAA 274
           |||||||||||||||||||||||||||||
Db   29 AAGAGCAAGAATGGGGGATTCACATCTAA 1

RESULT 3
AX330137/c
LOCUS   AX330137   145831 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION   Sequence 646 from Patent WO0194629.
ACCESSION   AX330137
VERSION   AX330137.1 GI:18103115
KEYWORDS
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
                Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE   Cancer gene determination and therapeutic screening using signature
                gene sets
  JOURNAL   Patent: WO 0194629-A 646 13-DEC-2001;
                Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
     source   1..145831
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   11.3%;   Score 98; DB 6; Length 145831;
Best Local Similarity   100.0%; Pred. No. 6.4e-46;
Matches 98;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGAGGAGGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGAGGAGGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
           ||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 4
AX334634/c
LOCUS   AX334634   145831 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION   Sequence 5143 from Patent WO0194629.
ACCESSION   AX334634
VERSION   AX334634.1 GI:18125353
KEYWORDS
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
                Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE   Cancer gene determination and therapeutic screening using signature
                gene sets
  JOURNAL   Patent: WO 0194629-A 5143 13-DEC-2001;
                Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
     source   1..145831
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   11.3%;   Score 98; DB 6; Length 145831;
Best Local Similarity   100.0%; Pred. No. 6.4e-46;
Matches 98;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGAGGAGGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGAGGAGGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
           ||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 5
AX336416/c
LOCUS   AX336416   145831 bp   DNA   linear   PAT 09-JAN-2002
DEFINITION   Sequence 6925 from Patent WO0194629.
ACCESSION   AX336416
VERSION   AX336416.1 GI:18127135
KEYWORDS
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
                Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE   Cancer gene determination and therapeutic screening using signature
                gene sets
  JOURNAL   Patent: WO 0194629-A 6925 13-DEC-2001;
                Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
     source   1..145831
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   11.3%;   Score 98; DB 6; Length 145831;
Best Local Similarity   100.0%; Pred. No. 6.4e-46;
Matches 98;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGAGGAGGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGAGGAGGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
           ||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 6

```
Qy       87 CGCCCGGCCTCCGGCCTCCCAAAGTGCTGGGGATTACAGGC 124
            |||||||||||||||||||||||||||||||||||||||||
Db    38540 CGCCCGGCCTCCGGCCTCCCAAAGTGCTGGGGATTACAGGC 38503
```

LOCUS   AX337041
DEFINITION Sequence 7550 from Patent WO0194629.
ACCESSION   AX337041
VERSION   AX337041.1 GI:18127760
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE   Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 7550 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
     source   1..145831
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   11.3%; Score 98; DB 6; Length 145831;
Best Local Similarity   100.0%; Pred. No. 6.4e-46;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       27 AGTAGAGATGGGGTTTCACCCTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    38600 AGTAGAGATGGGGTTTCACCCTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy       87 CGCCCGGCCTCCGGCCTCCCAAAGTGCTGGGGATTACAGGC 124
            |||||||||||||||||||||||||||||||||||||||||
Db    38540 CGCCCGGCCTCCGGCCTCCCAAAGTGCTGGGGATTACAGGC 38503
```

RESULT 7
HUAF001548/c

LOCUS   HUAF001548   145831 bp   DNA   linear   PRI 10-JAN-2000
DEFINITION Human Chromosome 16 BAC clone CIT987SK-A-815A9, complete sequence.
ACCESSION   AF001548
VERSION   AF001548.1 GI:2104552
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 145831)
  AUTHORS   Loftus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R.,
            Fuhrmann,J., Mason,T., Crosby,M.L., Barnstead,M., Cronin,L.,
            Deslattes Mays,A., Cao,Y., Xu,R.X., Kang,H.L., Mitchell,S.,
            Eichler,E.E., Harris,P.C., Venter,J.C. and Adams,M.D.
  TITLE   Genome duplications and other features in 12 Mb of DNA sequence
            from human chromosome 16p and 16q
  JOURNAL   Genomics 60 (3), 295-308 (1999)
  MEDLINE   99425270
   PUBMED   10493829
REFERENCE   2 (bases 1 to 145831)
  AUTHORS   Adams,M.D.
  TITLE   Direct Submission
  JOURNAL   Submitted (29-APR-1997) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
REFERENCE   3 (bases 1 to 145831)
  AUTHORS   Adams,M.D. and Loftus,B.J.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-AUG-1998) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
COMMENT   BAC clone CIT987SK-A-815A9 is located in chromosome 16. Genes
            were identified by a combination of five methods: XGRAIL (available
            by anonymous ftp from arthur.epm.ornl.gov), Genefinder (available
            by anonymous ftp from colin@u.washington.edu), GENSCAN (available
            using the e-mail server at genscan@gnomic.stanford.edu), searches
            of the EST database at TIGR (http://www.tigr.org/tdb/hcd/hcd.html)
            and searches against a peptide database. Repeats were identified
            using RepeatMasker (Smit, A.F.A. and Green, P. unpublished,
            http://ftp.genome.washington.edu/rm/RepeatMasker.html).
FEATURES   Location/Qualifiers
     source   1..145831
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /clone="A-815A9"
                     /map="16p13.11"
     gene   <9317..129591
                     /gene="A-815A9.1"
     mRNA   join(<9317..9473,34041..34068,45996..46098,50251..50343,
                     53885..53948,56552..56650,61016..61159,68837..68932,
                     72070..72188,73000..73152,74728..74901,76214..76387,
                     79220..79334,82397..82590,84560..84681,84785..84854,
                     84998..85158,87491..87599,90837..90968,91059..91265,
                     92540..92677,94040..94163,95108..95279,97150..97362,
                     100020..100164,105674..105880,107736..107840,
                     107929..108081,108319..108567,111094..111306,
                     111677..111889,112416..112577,113015..113143,
                     113420..113508,114290..114413,115380..115588,
                     117456..117564,117647..117819,128605..129591)
                     /gene="A-815A9.1"
     CDS   join(<9317..9473,34041..34068,45996..46098,50251..50343,
                     53885..53948,56552..56650,61016..61159,68837..68932,
                     72070..72188,73000..73152,74728..74901,76214..76387,
                     79220..79334,82397..82590,84560..84681,84785..84854,
                     84998..85158,87491..87599,90837..90968,91059..91265,
                     92540..92677,94040..94163,95108..95279,97150..97362,
                     100020..100164,105674..105880,107736..107840,
                     107929..108081,108319..108567,111094..111306,
                     111677..111889,112416..112577,113015..113143,
                     113420..113508,114290..114413,115380..115588,
                     117456..117564,117647..117819,128605..129591)
                     /gene="A-815A9.1"
                     /codon_start=1
                     /product="Myosin heavy chain (MYH11) (5'partial)"
                     /protein_id="AAC31665.1"
                     /db_xref="GI:2104553"
                     /translation="TYSGLFCVVVNPYKHLPIYSEKIVDMYKGKKRHEMPPHIYAIAD
                     LQANPLEAFGNAKTVKNDNSSRFGKFIRINFDVTGYIVGANIETVLLEKSRAIRQA
                     RDERTFHIFYYMIAGAKERMRSDLLLEGPNNYTFLSNGFVPIPAAQDDEMFQETVRAM
                     AIMGFSEEEQLSILKVVSSVLQLGNIVFKKERNTDQASMPDNTAAQKVCHLMGINVTD
                     FTRSILTPRIKVGRDVVQKAQTKEQADFAVEALAKATYERLFRWILTRVNKALDKTHR
                     QGASPLGILDIAGFEIFEVNSFEQLCINYTNEKLQQLFNFHMFILEQEEYQREGIEWN
                     FIDFGLDLQPCIELIERPNNPPGVLALLDEECWFPKATDKSFVEKLCTEQGSHPKFQK
                     PKQLKDKTEFSIIHYAGKVDYNASAWLTKNMDPLNDNVTSLLNASSDKFVADIWKDVD
                     RIVGLDQMAKMTESSLPSASKTKKGMFRTVGQLYKEQLGKLMTTLRNTNPNFVRCIIP
                     NHEKRSGKLDAPLVLEQLRCNGVLEGIRICRQGFPNRIVFQEFRQRYEILAANAIPKG
                     FMDGKQACILMIKALELDPNLYRIGQSKIFFRTVGQLAHLEEERDLKITDVIMAFQMC
                     RGYLARKAFAKRQQQLTAMKVIQRNCAAYLKLRNWQWWRLFTKVKPLLQVTRQEEMQ
                     AKEDELQKTKERQQKAENELKELEQKISQLTEKNLLQEQLQAETELYAEAEEMRVRL
                     AAKKQELEEILHEMEARLEEEEDRGQLQARKKKMAQQMLDLEEQLEEEARQKLQL
                     EKVTAEAKIKLEDEILVMDDQNNKLSKERKLLEERISDLTTNLAEEEKNLKTLK
                     NKHESMISELEVRLKEEKSRQELRKLRKLEGDASDFHEQLADLQQIALELKMQLAK
                     KEEELQAALARLDDEIAQKNNALKKIRELEGHISDLQEDLDSERAARNKAEKQKRDLG
                     AVEELTQLEQFKRAKANLDKNKQTLEKENADLAGELRVLGQAKQEVEHKKKLEAQV
                     BELRALKTELEDTLDSTATQQELRAKREQEVTVLKKALDEETRSHEAQVQEMRQKHAQ
                     QELQSKCSSDGERAELNDIKVHKLQNEVESVTGMLNEAEGKAIKLARDVASLSSQLQD
                     TQELLQEETRQKLNVSTKLRQLEERNSLQDQLDEEMEAKQNLERHISTLNIQLSDSK
                     KKLQDFASTVEALEEGKKRFQKEIENLTQQYEEKAAAYDKLEKTKNRLQQELDDLVVD
                     LDNQRQLVSNLEKKQKKFDQLLAEEKNISSKYADERDRAEAEAREKETKALSLARALB"

```
STS    EALEAKEELERTNKMLKAEMEDLVSSKDDVGKNVHELEKSKRALETQMEEMKTQLEEL
       EDELQATEEDAKYLRLEVNMQAQFQOFERDLQARDEQN2ERRROLQRQLHEYBELRDER
       KQRALAAAKKKLEGDLKDLRIQADSAIRGREEAIKQLRKLQAOMKDFORELEDARAS
       RDEIPATAXENEKKAXSLEADLMQLQEDLAAAFARARKQADLEKEELASSLSGRN
       ALQDEKRRLEARIAQLEEELEEEQGNMEAMSDRVRKATQQEAEQLSNELATERSTAQKN
       ESAQOLEROMKELSKLLHEMEGAVRSKFKFSTIAALEAKIAQLEEQVEQBARAKEQAAT
       KSIXOKDKLKLELLQVEDERKKAROYKEQAEKQNARVKQLKRQLEEAEESQRINAN
       RRKLQRELDEATESNEAMGREVNALKSKLRRGNETSFVPSRSGGRRVTENADGSEEE
       TDTRDADFNGTKASE"
       Taiwan;
       11958..12213
       /gene="A-815A9.1"
       /db_xref="dbSTS:12211"

STS    30562..30695
       /gene="A-815A9.1"
       /db_xref="dbSTS:G17920"

STS    43852..43993
       /gene="A-815A9.1"
       /db_xref="dbSTS:G13936"

STS    ..........
       /gene="A-815A9.1"
       /db_xref="dbSTS:Z16518"

STS    6285..6480
       /gene="A-815A9.1"
       /db_xref="dbSTS:G23513"

STS    12562..126043
       /gene="A-815A9.1"
       /db_xref="dbSTS:252849"

STS    76508..76638
       /gene="A-815A9.1"
       /db_xref="dbSTS:G17876"

STS    10231..10838
       /gene="A-815A9.1"
       /db_xref="dbSTS:197033"

STS    12627..128736
       /gene="A-815A9.1"
       /db_xref="dbSTS:G13593"

ORIGIN

Query Match     11.3%;  Score 98;  DB 9;  Length 145831;
Best Local Similarity   100.0%;  Pred. No. 6.4e-46;
Matches  98;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      27 AGTAGAGAATGGGGTTCACCCGGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC  86
           ||||||||| ||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGGGGTTCACCCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy      87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
           ||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 38503
```

```
RESULT 8
LOCUS       BS000050           202014 bp    DNA     linear   PRI 07-OCT-2003
DEFINITION  Pan troglodytes chromosome 22 clone:RP43-038C12, map 22, complete
            sequences.
ACCESSION   BS000050
VERSION     BS000050.1  GI:37537317
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   The Chimpanzee Chromosome 22 Sequencing Consortium.
  TITLE     DNA sequence of chimpanzee chromosome 22 and its evolutionary
            implications
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202014)
  AUTHORS   Hattori,M., Toyoda,A., Watanabe,H., Taylor,T.D., Kuroki,Y.,
            Fujiyama,A. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp-gsc.riken.go.jp/,
```

```
COMMENT     Tel:81-45-503-9111, Fax:81-45-503-9170)
            ------------------------------------------------------------
            *Chinese National Human Genome Center at Shanghai, Shanghai, China;
            *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
            of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
            Center, Daejeon, Korea;
            *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
            *National Institute of Genetics, Mishima, Japan;
            *National Yang Ming University Genome Research Center, Taipei,
            Taiwan;
            ------------------------------------------------------------
            *RIKEN Genomic Sciences Center, Yokohama, Japan.
            ------------------------------------------------------------
            Center: RIKEN Genomic Sciences Center
            Center code: RIKEN
            Web site: http://hgp.gsc.riken.go.jp/
            Contact: hattori@gsc.riken.go.jp
            ------------------------------------------------------------
            Center project name:The Chimpanzee Chromosome 22 Sequencing Project
            Center clone name: RP43-038C12
            ------------------------------------------------------------
                           Summary Statistics
            Sequencing vector: pUC18,pUC13,pTZ19R; 100% of reads Chemistry:
            Dye-terminator Big Dye and ET; 100% of reads Assembly program:
            Phrap; version 0.990329
            Consensus quality:   200,900 bases at least Q40
            Consensus quality:       633 bases at least Q30
            Consensus quality:        83 bases at least Q20
            ------------------------------------------------------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30;
            an attempt was made to resolve all sequencing problems, such as
            compressions and repeats; all regions were covered by at one
            plasmid
            subclone or more than one M13 subclone;
            and the assembly was confirmed by restriction digest.
            Source information:
            The RPCI-43 chimpanzee BAC library was prepared from DNA isolated
            from the blood of a single male chimpanzee using published
            protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the
            chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
            Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
            Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
            characteristics are described at
            http://www.chori.org/bacpac/chimpanzee251.htm.
            The clone may be obtained from Pieter J. de Jong and coworkers
            (http://www.chori.org/bacpac).
            VECTOR: pBACe3.6
            The CHORI-251 chimpanzee BAC library was prepared from DNA isolated
            from the blood of a single male chimpanzee using published
            protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the
            chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
            Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
            Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
            characteristics are described at
            http://www.chori.org/bacpac/chimpanzee251.htm.
            The clone may be obtained from Pieter J. de Jong and coworkers
            (http://www.chori.org/bacpac).
            VECTOR: pTARBAC2.1
            The PTB1 chimpanzee BAC library was prepared from DNA isolated from
            cultured cells established from the blood of a single male
            chimpanzee.
            Clones may be obtained from Asao Fujiyama and co-workers
            (http://www.gsc.riken.go.jp).
            VECTOR: pKS145
            The PTF22 chimpanzee Fosmid library was prepared from DNA isolated
            from cultured cells established from the blood of a single male
            chimpanzee.
            Clones may be obtained from Asao Fujiyama and co-workers
            (http://www.gsc.riken.go.jp).
            VECTOR: pKS143
            ------------------------------------------------------------
```

Sequence Quality Assessment:
This entry has been annotated with sequence
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in
10,000 bp.

Neighboring clones: RP43-069C10(left) and RP43-089J14(right).

FEATURES
    source   Location/Qualifiers
                    1..202014
                    /organism="Pan troglodytes"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9598"
                    /chromosome="22"
                    /clone="RP43-038C12"
                    /clone_lib="RPCI-43 chimpanzee BAC"

ORIGIN
    Query Match   11.3%; Score 98; DB 9; Length 202014;
    Best Local Similarity 100.0%; Pred. No. 6.4e-46;
    Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    Qy   27 AGTAGAGGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db   91024 AGTAGAGGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 90965

    Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
               ||||||||||||||||||||||||||||||||||||||
    Db   90964 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 90927

RESULT 9
LOCUS   AC021026/c
DEFINITION AC021026/c
            Homo sapiens 3 BAC RP11-85A3 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC021026
VERSION   AC021026.20 GI:22002087
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

LOCUS   AC021026   75555 bp   DNA   linear   PRI 29-JUL-2002

REFERENCE   1 (bases 1 to 75555)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alebrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,B., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.I., Byrd,N.C.,
            Carron,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Louiseed,H., Lozado,R.J., Lu,X., Lucier,R., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,B., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,Z., Mitchell,T., Mohabbat,K.,
            Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
            Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
            Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,

(continued)

            Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
            Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
            Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
            Scott,C., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
            Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
            Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
            Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
            Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
            Warren,R., Washington,C., Watlington,S., Williams,G.,
            Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
            Gibbs,R.
            Wu,Y.P., Zhou,J., Zorrilla,S., Naylor,S.L., Naylor,S.L., Weinstock,G. and

  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 75555)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-JAN-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 75555)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (14-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4 (bases 1 to 75555)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (29-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Jul 29, 2002 this sequence version replaced gi:2174731.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
            STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
            Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
            Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as Low Coverage.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht
            ml.

FEATURES   Location/Qualifiers
    source   1..75555
                    /organism="Homo sapiens"

```
                 /mol_type="genomic DNA"
                 /db_xref="taxon:9606"
                 /chromosome="3"
                 /clone="RP11-85A3"
repeat_region    complement(1..177)
                 /rpt_family="AluSq"
repeat_region    complement(190..474)
                 /rpt_family="AluSq"
repeat_region    complement(475..650)
                 /rpt_family="AluJo"
repeat_region    1319..1534
                 /rpt_family="MIR"
repeat_region    i694..2002
                 /rpt_family="HAL1"
repeat_region    complement(2305..2600)
                 /rpt_family="AluJb"
repeat_region    3279..3575
                 /rpt_family="AluSq"
repeat_region    4477..4520
                 /rpt_family="AT_rich"
repeat_region    4569..4861
                 /rpt_family="AluJb"
repeat_region    complement(5022..5248)
                 /rpt_family="MIR"
repeat_region    complement(6674..6974)
                 /rpt_family="AluSx"
repeat_region    7198..7367
                 /rpt_family="AluJo"
repeat_region    7372..7538
                 /rpt_family="AluJo"
repeat_region    7560..7628
                 /rpt_family="(TA)n"
repeat_region    7644..7685
                 /rpt_family="Alu"
repeat_region    7687..7723
                 /rpt_family="(TA)n"
repeat_region    7792..7842
                 /rpt_family="(TA)n"
repeat_region    10037..10393
                 /rpt_family="AluJo"
repeat_region    complement(10400..10697)
                 /rpt_family="MER47A"
repeat_region    complement(7853..7894)
                 /rpt_family="(TAGA)n"
repeat_region    7904..7987
                 /rpt_family="Alu"
repeat_region    complement(10897..11198)
                 /rpt_family="AluJb"
repeat_region    11230..11571
                 /rpt_family="AluJb"
repeat_region    11572..11871
                 /rpt_family="LIMB4"
repeat_region    11872..12036
                 /rpt_family="AluJo"
repeat_region    12037..12349
                 /rpt_family="LIMB4"
repeat_region    12350..12796
                 /rpt_family="AluSq"
repeat_region    complement(12797..12831)
                 /rpt_family="LIMB4"
repeat_region    12832..12954
                 /rpt_family="U2"
repeat_region    12955..13256
                 /rpt_family="LIMB4"
repeat_region    13257..13377
                 /rpt_family="AluY"
repeat_region    complement(13378..13608)
                 /rpt_family="LIMB4"
repeat_region    13809..13974
                 /rpt_family="MSTB"
repeat_region    13989..14054
                 /rpt_family="LIMB4"
repeat_region    14062..14090
                 /rpt_family="LIMB8"
repeat_region    complement(14254..14312)
                 /rpt_family="(A)n"
repeat_region    14470..14765
                 /rpt_family="L2"
repeat_region    14887..14916
                 /rpt_family="AluSx"
repeat_region    15711..15899
                 /rpt_family="AT_rich"
misc_feature     15712..15897
                 /note="size confirmed by restriction map, formally
                 approved."
                 /function="unresolved tandem repeat"


Query Match    10.3%;  Score 89;  DB 9;  Length 75555;
Best Local Similarity  100.0%;   Pred. No. 1.3e-40;
Matches  89;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    36 GGGGTTTCACCGTGGTTAGGCCAGGATGGGTGTTCGGAACTTCCTGACCTGTGATCGGCCGCCCGCCCT 95
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 13063 GGGGTTTCACCGTGGTTAGGCCAGGATGGGTGTTCGGAACTTCCTGACCTGTGATCGGCCGCCCGCCCT 13004

Qy    96 CGGCCTCCCAAAAGTGCTGGGATTACAGGC 124
         ||||||||||||||||||||||||||||||
Db 13003 CGGCCTCCCAAAAGTGCTGGGATTACAGGC 12975

RESULT 10
AC004537
LOCUS       AC004537               88872 bp    DNA     linear   PRI 08-OCT-2003
DEFINITION  Homo sapiens PAC clone RP5-872F7 from 7, complete sequence.
ACCESSION   AC004537
VERSION     AC004537.1  GI:3041854
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88872)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   MEDLINE   99063792
   PUBMED    9847074
REFERENCE   2  (bases 1 to 88872)
  AUTHORS   Burkhart,J., Dauphin,S. and Elliott,G.
  TITLE     The sequence of Homo sapiens PAC clone RP5-872F7
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 88872)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-APR-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   4  (bases 1 to 88872)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 88872)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     -------------- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
            Center project name: H_DJ0872F07
```

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

SOURCE INFORMATION:
This clone was derived from human PAC library RPCI-4, prepared by
Pieter de Jong and coworkers at http://www.chori.org using the
method described by Ioannou et al., Nature Genetics 6:84-9 (1994).
The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc.
(http://www.genomesystems.com) or Research Genetics, Inc.
(http://www.resgen.com); or from Pieter de Jong.
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RPS-1100F23, 200 bp overlap.
Actual start of this clone is at base position 111440 of
RP5-1100F23 actual end is at 88872 of RP5-872F7.

```
FEATURES             Location/Qualifiers
     source          1..88872
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP5-872F7"
                     /clone_lib="RPCI-4"
     repeat_region   730..1048
                     /rpt_family="Alu"
     repeat_region   2125..2155
                     /rpt_family="(TG)n"
     repeat_region   4730..4827
                     /rpt_family="L1"
     repeat_region   5487..6038
                     /rpt_family="L1"
     repeat_region   6013..6188
                     /rpt_family="AcHobo"
     repeat_region   6189..6486
                     /rpt_family="AcHobo"
     repeat_region   5487..6675
                     /rpt_family="AcHobo"
     repeat_region   6676..6991
                     /rpt_family="AcHobo"
     repeat_region   6992..7454
                     /rpt_family="Alu"
     repeat_region   7455..7759
                     /rpt_family="AcHobo"
     repeat_region   7760..8206
                     /rpt_family="Alu"
     repeat_region   8209..8674
                     /rpt_family="AcHobo"
     repeat_region   8676..8778
                     /rpt_family="MaLR"
     repeat_region   8973..9747
                     /rpt_family="AcHobo"
     repeat_region   9751..9789
                     /rpt_family="L1"
     repeat_region   9856..10618
                     /rpt_family="L1"
     repeat_region   11330..11364
                     /rpt_family="ERV1"
     repeat_region   11378..13569
                     /rpt_family="AT_rich"
     repeat_region   13582..13808
                     /rpt_family="MIR"
     repeat_region   15324..16268
                     /rpt_family="MIR"
     repeat_region   16428..16709
                     /rpt_family="L1"
     repeat_region   16710..17023
                     /rpt_family="ERVL"
     repeat_region   17376..17684
                     /rpt_family="MaLR"
     repeat_region   17924..18068
                     /rpt_family="ERVL"
     repeat_region   18239..18307
                     /rpt_family="ERVL"
     repeat_region   18893..19345
                     /rpt_family="ERV1"
     repeat_region   19373..19528
                     /rpt_family="ERVL"
     repeat_region   19535..19671
                     /rpt_family="L2"
     repeat_region   19673..19961
                     /rpt_family="(TA)n"
     repeat_region   19962..20504
                     /rpt_family="MER2_type"
     repeat_region   20910..21757
                     /rpt_family="L2"
     repeat_region   21861..21888
                     /rpt_family="ERVL"
     repeat_region   22266..22574
                     /rpt_family="(A)n"
     repeat_region   23531..23563
                     /rpt_family="L1"
     repeat_region   24301..24416
                     /rpt_family="AT_rich"
     repeat_region   25598..25397
                     /rpt_family="L2"
     repeat_region   27766..28004
                     /rpt_family="MER1_type"
     repeat_region   28033..28091
                     /rpt_family="MER1_type"
     repeat_region   28498..28647
                     /rpt_family="L2"
     repeat_region   28065..28868
                     /rpt_family="L2"
     repeat_region   28384..28500
                     /rpt_family="MIR"
     repeat_region   28707..29021
                     /rpt_family="L1"
     repeat_region   29978..30276
                     /rpt_family="L1"
     repeat_region   30494..30671
                     /rpt_family="Alu"
     repeat_region   31539..32022
                     /rpt_family="MIR"
     repeat_region   32170..32274
                     /rpt_family="L1"
     repeat_region   33275..32435
                     /rpt_family="L1"
     repeat_region   33275..32435
                     /rpt_family="GA-rich"
```

misc_feature   33451..34194
    /note="CpG island (%GC=61.2, o/e=0.87, #CpGs=73)"
repeat_region 33669..33695
    /rpt_family="(T)n"
gene   33583..58210
    /gene="ING3"
mRNA   join(33683..33844,34084..34155,36224..36324,38479..38544,
               47662..47758,49546..49617,50449..50568,50854..51011,
               51931..52124,53608..53800,56125..56163,57648..58210)
               /gene="ING3"
mRNA   join(33699..33844,34084..34155,36224..36324,38479..38544,

Query Match
Best Local Similarity   10.3%;
Matches   89; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   36 GGGGTTTCACCGTGTTGTAGCCAGGATGGTCTCGAACTCCTGGTGATCCGCCCGCCT 95
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42756 GGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTCGAACTCCTGATCCGCCCGCCT 42815

Qy   96 CGGGCCTTCCAAAGTGCTGGGATTACAGGC 124
           ||||||||||||||||||||||||||||||
Db   42816 CGGGCCTTCCAAAGTGCTGGGATTACAGGC 42844

RESULT 11
AC004840
LOCUS   AC004840   149397 bp   DNA   linear   PRI 09-MAR-2002
DEFINITION Homo sapiens PAC clone RP4-607J2 from 7, complete sequence.
ACCESSION   AC004840
VERSION   AC004840.4 GI:19310320
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 149397)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE   Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2 (bases 1 to 149397)
  AUTHORS   Edwards,J., Wollam,A., Walker,C., Gregory,S., Hawkins,M. and
            Fiedler,T.
  TITLE   The sequence of Homo sapiens PAC clone RP4-607J2
  JOURNAL   Unpublished (2001)
REFERENCE   3 (bases 1 to 149397)
  AUTHORS   Waterston,R.H.
  TITLE   Direct Submission
  JOURNAL   Submitted (07-OCT-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   4 (bases 1 to 149397)
  AUTHORS   Waterston,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5 (bases 1 to 149397)
  AUTHORS   Waterston,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (09-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT   On Mar 9, 2002 this sequence version replaced gi:1071623.
            --------------------- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            --------------------- Summary Statistics
            Center project name: H_DJ0607J02

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/GTB/CHR7, send
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
This clone was derived from human PAC library RPCI-4, prepared by
Pieter de Jong and coworkers at http://www.chori.org using the
method described by Ioannou et al., Nature Genetics 6:84-9 (1994).
The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc.
(http://www.genomesystems.com) or Research Genetics, Inc.
(http://www.resgen.com); or from Pieter de Jong.
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RP5-1125K23, 200 bp overlap.
The clone sequenced to the left is RP5-1125K3, 200 bp overlap.
Actual start of this clone is at base position 1 of RP4-607J2.
actual end is at base position 149397 of RP4-607J2.

There are polymorphic base changes between RP4-607J2 and
CTD-2312I3. The sequence RP4-607J2 contains an unresolvable tandem
repeat from base position 21064 to 22995 whose sequence fidelity
can not be guaranteed. Digest information suggests that there are
82 bases missing in the sequence.

FEATURES   Location/Qualifiers
     source   1..149397
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP4-607J2"
                     /clone_lib="RPCI-4"
     repeat_region   514..578
                     /rpt_family="(TCCA)n"
     repeat_region   819..912
                     /rpt_family="G-rich"
     repeat_region   1518..1700
                     /rpt_family="MER1_type"
     repeat_region   2435..2511
                     /rpt_family="MIR"
     repeat_region   3423..3491
                     /rpt_family="MIR"
     repeat_region   4582..4669
                     /rpt_family="MER1_type"
     repeat_region   5780..5847
                     /rpt_family="MIR"
     repeat_region   6966..7279
                     /rpt_family="MaLR"
     repeat_region   7280..7591
                     /rpt_family="Alu"
     repeat_region   7592..7680
                     /rpt_family="MIR"
     repeat_region   8860..8880
                     /rpt_family="MaLR"

```
repeat_region   9307..9359     /rpt_family="GC_rich"
repeat_region   9378..9535     /rpt_family="MIR"
repeat_region   10352..10420   /rpt_family="MIR"
repeat_region   11955..11997   /rpt_family="MIR"
repeat_region   12114..12226   /rpt_family="CR1"
misc_feature                   /note="match to EST AW882247 (NID:g8044257)"
repeat_region   12291..12575   /rpt_family="Alu"
misc_feature    1298..12312    /note="match to EST BP511578 (NID:g11594876)"
misc_feature    12317..12368   /note="match to EST BG149317 (NID:g12661347)"
repeat_region   12577..12911   /rpt_family="Alu"
repeat_region   12747..13220   /rpt_family="Alu"
misc_feature    13489..13917   /note="match to EST AA210922 (NID:g1809818) zq89g01.s1"
misc_feature    16636..16651   /note="match to EST AA211096 (NID:g1809594) zq89g01.r1"
repeat_region   16652..16966   /rpt_family="L2"
repeat_region   16967..17022   /rpt_family="Alu"
repeat_region   18362..18664   /rpt_family="L2"
misc_feature    18805..19093   /note="match to EST AA774302 (NID:g2826111) ac37b12.s1"
repeat_region   19689..19938   /rpt_family="(ACCTG)n"
repeat_region   20071..20183   /rpt_family="(CAGAGA)n"
repeat_region   21010..21167   /rpt_family="(TGGA)n"
repeat_region   21168..21350   /rpt_family="Alu"
repeat_region   21395..21437   /rpt_family="(GCTG)n"
repeat_region   21459..21637   /rpt_family="(TGGA)n"
repeat_region   21653..21754   /rpt_family="(TGGG)n"
repeat_region   21781..22322   /rpt_family="(TGGA)n"
repeat_region   22359..22717   /rpt_family="(TGGA)n"
repeat_region   23315..23406   /rpt_family="MIR"
repeat_region   24007..24310   /rpt_family="Alu"
repeat_region   25412..25605   /rpt_family="Alu"
repeat_region   26559..26865   /rpt_family="MIR"
repeat_region   26882..27183   /rpt_family="Alu"
repeat_region   27185..27246   /rpt_family="Alu"
repeat_region   27247..27548   /rpt_family="Alu"
repeat_region   27549..27567   /rpt_family="Alu"
repeat_region   27568..27855   /rpt_family="MER1_type"
repeat_region   27856..27882   /rpt_family="Alu"
repeat_region   27883..28182   /rpt_family="MER1_type"
repeat_region                  /rpt_family="Alu"
```

```
repeat_region   28183..28251   /rpt_family="MER1_type"
repeat_region   28252..28543   /rpt_family="Alu"
repeat_region   28544..28614   /rpt_family="Alu"
repeat_region   28615..28740   /rpt_family="MER1_type"
repeat_region                  /rpt_family="(TG)n"
```

```
Query Match
Best Local Similarity   10.3%;   Score 89;  DB 9;  Length 149397;
Matches   89;  Conservative   0;  Pred. No. 1.3e-40;
          89;                 0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     36 GGGGTTCACCCGTTGTTAGCCAGAGATGGTCTCGAAACTTCTTGACTTCGTGATCCCGGCCT 95
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 111428 GGGGTTCACCCGTTGTTAGCCAGAGATGGTCTCGAAACTTCTTGACTTCGTGATCCCGGCCT 111487

Qy     96 CGGGCCTCCCAAAGCTGCTGGGATTACAGGC 124
          |||||||||||||||||||||||||||||||
Db 111488 CGGGCCTCCCAAAGCTGCTGGGATTACAGGC 111516
```

TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 152005)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (12-JAN-2000) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 152005)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (08-JUN-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 152005)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (25-JUN-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 152005)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (26-JUL-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 6 (bases 1 to 152005)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (15-MAR-2003) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 7 (bases 1 to 152005)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (14-MAR-2003) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
           On Jun 25, 2002 this sequence version replaced gi:21328480.
COMMENT   INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
           gc-help@bcm.tmc.edu

           CLONE LENGTH: This sequence does not necessarily represent the
           entire insert of this clone. Overlapping regions of clones are only
           sequenced and submitted once, so the sequence for the remainder of
           the insert may be found in the record for the adjacent clones.
           Overlapping clones are noted at the beginning and end of the
           Features listing.

           ANNOTATION OF FEATURES:
           STSs are identified using ePCR (Genome Res. 7:541-550) searches
           of a local database that includes entries from dbSTS, GDB, and
           local mapping efforts.

           Repeats are identified using RepeatMasker (A. Smit and P. Green,
           unpublished.) for Human and Mouse sequences.

           Genes and Region of sequence similarity are identified by BLAST
           (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
           EST and cDNA sequences. Genes demonstrate at least two exons
           flanked by consensus splice sites that maintained sequence
           continuity across the splice junctions. Sequences that are not
           identical matches are annotated as similar.

           SEQUENCING READ COVERAGE:Sequencing is completed to a minimum of 2
           standard of double strand coverage with a minimum of 2 clones and 2
           reads with no ambiguities or 2 chemistries with a minimum of 2
           clones and 3 reads with no ambiguities. If the sequence quality for
           a region does not meet this standard, it will be indicated in the
           annotation as Low Coverage.

           QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
           standards - estimated error rate less than 1 per 10,000 bases.
           Reports of lowest quality individual bases and measures of base
           quality are listed below. Description of the metrics can be found
           at URL:
           http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht
           ml.

FEATURES   Location/Qualifiers
     source   1..152005
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-79L9"
     misc_feature   1..2004
                     /note="overlaps bases 109101..111104 of clone AC117389"
                     /function="clone overlap"
     repeat_region   324..628
                     /rpt_family="AluSx"
     repeat_region   complement(633..930)
                     /rpt_family="AluSx"
     repeat_region   1359..1647
                     /rpt_family="AluSx"
     repeat_region   2377..2421
                     /rpt_family="AluSg"
     repeat_region   complement(3409..3712)
                     /rpt_family="AT_rich"
     repeat_region   3873..4041
                     /rpt_family="AluJb"
     repeat_region   5445..5743
                     /rpt_family="MIR"
     repeat_region   5798..6102
                     /rpt_family="AluSq"
     repeat_region   6179..6210
                     /rpt_family="AT_rich"
     repeat_region   complement(6403..6703)
                     /rpt_family="AluSg"
     repeat_region   6791..6878
                     /rpt_family="T-rich"
     repeat_region   6892..7006
                     /rpt_family="AluJb"
     repeat_region   7007..7036
                     /rpt_family="(TAA)n"
     repeat_region   complement(7599..7682)
                     /rpt_family="LIM4"
     repeat_region   7037..7203
                     /rpt_family="AluJb"
     repeat_region   7326..7440
                     /rpt_family="AluJb"
     repeat_region   complement(7717..7946)
                     /rpt_family="AluSx"
     repeat_region   7443..7470
                     /rpt_family="(TAA)n"
     repeat_region   complement(7947..8149)
                     /rpt_family="LIM4"
     repeat_region   complement(8413..8509)
                     /rpt_family="AluJo/FRAM"
     repeat_region   complement(8536..8794)
                     /rpt_family="AluSp"
     repeat_region   complement(8877..9170)
                     /rpt_family="AluSp"
     repeat_region   complement(9388..9427)
                     /rpt_family="HAl1"
     repeat_region   9428..9726
                     /rpt_family="AluSq"
     repeat_region   complement(9727..9760)
                     /rpt_family="AluSx"
     repeat_region   complement(9770..10778)
                     /rpt_family="L1MB8"
     repeat_region   complement(10761..11041)
                     /rpt_family="L1MB8"
     repeat_region   complement(11032..11705)
                     /rpt_family="FLAM_C"

repeat_region   /rpt_family="L1M4"
                   11706..11728
                   /rpt_family="(TTTG)n"
repeat_region   complement(11729..12011)
                   /rpt_family="AluSx"

Query Match
Best Local Similarity   10.3%;   Score 89; DB 9; Length 152005;
Matches   89; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   36 GGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGAACTCGTGATCCGCCCGCCT 95
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 74493 GGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGAACTCGTGATCCGCCCGCCT 74434

Qy   96 CGGGCCTCCCAAAGTGCTGGGATTACAGGC 124
         ||||||||||||||||||||||||||||||
Db 74433 CGGGCCTCCCAAAGTGCTGGGATTACAGGC 74405

RESULT 13
AL683875/c
LOCUS   AL683875   184927 bp   DNA   HTG   02-APR-2003
DEFINITION Homo sapiens chromosome X clone RP11-446K8, 5 unordered pieces.
ACCESSION   AL683875
VERSION   AL683875.25 GI:29489394
KEYWORDS   HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 184927)
  AUTHORS   Lawlor,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On Apr 2, 2003 this sequence version replaced gi:28971596.
COMMENT   --------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------- Project Information
            Center project name: bA446K8
            --------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 86% of reads
            Chemistry: Dye-terminator Big Dye; 13% of reads
            Chemistry: Dye-terminator Big Dye; 1% of reads
            Consensus quality: 18278 bases at least Q40
            Consensus quality: 18293 bases at least Q30
            Consensus quality: 183854 bases at least Q20
            Assembly program: XGAP4; version 4.5
            Insert size: 184527; sum-of-contigs
            Insert size: 179006; 2.7% error; agarose-fp
            Quality coverage: 9.16x in Q20 bases; sum-of-contigs Quality
            coverage: 10.59x in Q20 bases; agarose-fp
            ----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            ----
            *   1   49340: contig of 49340 bp in length
            *   4941   49440: gap of 100 bp
            *   4941   49479: contig of 42039 bp in length
            *   91480   91579: gap of 100 bp
            *   91580   94862: contig of 3283 bp in length
            *   94863   94962: gap of 100 bp
            *   94963   119846: contig of 24884 bp in length
            * 119847   119946: gap of 100 bp in length
            * 119947   184927: contig of 64981 bp in length.

FEATURES   Location/Qualifiers
     source   1..184927
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /clone="RP11-446K8"
                     /clone_lib="RPCI-11.2"
     misc_feature   1..49540
                     /note="assembly_fragment:04719
                     clone_end:T7
                     vector_side:left"
     misc_feature   49441..91479
                     /note="assembly_fragment:00421"
     misc_feature   91580..94862
                     /note="assembly_fragment:01738"
     misc_feature   94963..119846
                     /note="assembly_fragment:04642"
     misc_feature   119947..184927
                     /note="assembly_fragment:04655
                     clone_end:SP6
                     vector_side:right"

ORIGIN
Query Match
Best Local Similarity   10.3%;   Score 89; DB 2; Length 184927;
Matches   89; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   36 GGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGAACTCGTGATCCGCCCGCCT 95
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 169677 GGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGAACTCGTGATCCGCCCGCCT 169618

Qy   96 CGGGCCTCCCAAAGTGCTGGGATTACAGGC 124
          ||||||||||||||||||||||||||||||
Db 169617 CGGGCCTCCCAAAGTGCTGGGATTACAGGC 169589

RESULT 14
AL683807/c
LOCUS   AL683807   189825 bp   DNA   linear   PRI 23-NOV-2002
DEFINITION Human DNA sequence from clone RP13-297E16 on chromosome X, complete
            sequence.
ACCESSION   AL683807
VERSION   AL683807.22 GI:25251452
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 189825)
  AUTHORS   Grafham,D.
  TITLE   Direct Submission
  JOURNAL   Submitted (22-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On Nov 24, 2002 this sequence version replaced gi:24474461.
COMMENT   --------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk

            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission with
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

FEATURES
     source   Location/Qualifiers
                     1. .189825
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /chromosome="X"
                     /db_xref="taxon:9606"
                     /clone="RP13-297B16"
                     /clone_lib="RPCI-13.2"

ORIGIN
     Query Match   10.3%; Score 89; DB 9; Length 189825;
     Best Local Similarity   100.0%; Pred. No. 1.3e-40;
     Matches   89; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   36 GGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATCCCGCCCGCCT   95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   88862 GGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATCCCGCCCGCCT 88803

Qy   96 CGGCCTCCCAAAGTGCTGGGATTACAGGC   124
              |||||||||||||||||||||||||||||
Db   88802 CGGCCTCCCAAAGTGCTGGGATTACAGGC 88774

RESULT 15
LOCUS   AL355146   57447 bp   DNA   linear   PRI 21-DEC-2000
DEFINITION   Human DNA sequence from clone RP11-31101 on chromosome 6, complete
               sequence.
ACCESSION   AL355146
VERSION   AL355146.13 GI:11967512
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 57447)
  AUTHORS   Tracey,A.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
               CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
               requests: clonerequest@sanger.ac.uk
COMMENT   On Dec 22, 2000 this sequence version replaced gi:11876034.
               During sequence assembly data is compared from overlapping clones.
               Where differences are found these are annotated as variations
               together with a note of the overlapping clone name. Note that the
               variation annotation may not be found in the sequence submission
               corresponding to the overlapping clone, as we submit sequences with
               only a small overlap as described above.
               This sequence has been finished according to sequence map criteria
               as follows. An attempt is made to resolve all sequencing problems,
               such as compressions and repeats, but not necessarily within known
               annotated repeat sequence elements. Where the sequence is
               ambiguous, there is an annotation using the 'unsure' feature key.
               The following abbreviations are used to associate primary accession
               numbers given in the feature table with their source databases:
               Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
               on the WORMPEP database can be found at

FEATURES
     source   Location/Qualifiers
                     1. .57447
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /chromosome="6"
                     /db_xref="taxon:9606"
                     /clone="RP11-31101"
                     /clone_lib="RPCI-11.2"

     repeat_region   1336. .1506
                     /note="L1M2 repeat: matches -668. .-492 of consensus"
     repeat_region   1516. .1540
                     /note="MER91 repeat: matches 100. .124 of consensus"
     repeat_region   1835. .2687
                     /note="L1 repeat: matches 3631. .4476 of consensus"
     repeat_region   2664. .2937
                     /note="L1M1 repeat: matches 5517. .5793 of consensus"
     repeat_region   2938. .2965
                     /note="14 copies a 2 mer ca 100% conserved"
     repeat_region   2967. .3380
                     /note="L1PA16 repeat: matches 5783. .6149 of consensus"
     misc_feature   complement(3594. .4086)
                     /note="match: GSS: Em:AQ451181"
     repeat_region   3898. .4140
                     /note="match: GSS: Em:AQ451181"
     misc_feature   complement(4026. .4542)
                     /note="match: GSS: Em:AQ720695"
     repeat_region   4247. .4545
                     /note="AluSx repeat: matches 7. .302 of consensus"
     repeat_region   4871. .5162
                     /note="AluSg repeat: matches 1. .292 of consensus"
     repeat_region   5207. .6253
                     /note="MIR repeat: matches 58. .245 of consensus"
     repeat_region   5563. .6744
                     /note="L1MA6 repeat: matches 5151. .6299 of consensus"
     repeat_region   6563. .6774
                     /note="L1MB2 repeat: matches 5812. .6002 of consensus"
     repeat_region   6774. .7192
                     /note="L1MB2 repeat: matches 5312. .5763 of consensus"
     repeat_region   7227. .7497
                     /note="L1ME2 repeat: matches 4962. .5246 of consensus"
     repeat_region   7505. .7628
                     /note="L1ME2 repeat: matches 5763. .6265 of consensus"
     repeat_region   7647. .7963
                     /note="L1MA8 repeat: matches 6141. .6265 of consensus"
     repeat_region   8755. .8938
                     /note="L1M4 repeat: matches 2670. .3000 of consensus"
     repeat_region   9625. .9854
                     /note="MIR repeat: matches 24. .262 of consensus"
     repeat_region   12052. .12222
                     /note="MLT1F repeat: matches 360. .533 of consensus"
     repeat_region   14872. .14936
                     /note="AluJr/FRAM repeat: matches 232. .293 of consensus"
     misc_feature   complement(15388. .15883)
                     /note="match: GSS: Em:AQ719416"
     repeat_region   15566. .15613
                     /note="24 copies a 2 mer ca 72% conserved"
     repeat_region   18406. .24551
                     /note="L1PA3 repeat: matches 3. .6146 of consensus"
     repeat_region   24575. .24903
                     /note="MER58B repeat: matches 1. .341 of consensus"
     repeat_region   27188. .27305

(other column, rotated page top material)

was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP13-297B16 is from the library RPCI-13.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-31101 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-31101 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP11-135M8 is at 57348 in this sequence.
The true left end of clone RP11-31101 is at 1 in this sequence. The
true left end of clone RP11-31101 is at 57447 in this sequence.

```
repeat_region                          /note="MIR repeat: matches 81. .205 of consensus"
repeat_region   27440. .27491          /note="L2 repeat: matches 2639. .2690 of consensus"
repeat_region   27635. .27743          /note="AluSx repeat: matches 2432. .2534 of consensus"
repeat_region   28606. .28672          /note="L2 repeat: matches 238. .302 of consensus"
misc_feature    28701. .28874          /note="ORSL repeat: matches 238. .302 of consensus"
misc_feature    28701. .28874  complement(28771. .28966) /note="match: GSS: Em:AQ629369"
repeat_region   28703. .28786          /note="match: GSS: Em:AQ629369"
repeat_region   28711. .28916          /note="AluJo/FRAM repeat: matches 199. .282 of consensus"
repeat_region   28793. .28834  complement(28771. .28966) /note="match: GSS: Em:AZ383385"
repeat_region   28837. .28916          /note="221 copies 2 mer ta 85% conserved"
repeat_region   29064. .29313          /note="40 copies 2 mer, at 76% conserved"
repeat_region   29314. .29607          /note="L1MA4 repeat: matches 4739. .4986 of consensus"
repeat_region   29608. .30913          /note="AluSx repeat: matches 1. .306 of consensus"
repeat_region   30946. .31112          /note="L1MA4 repeat: matches 4986. .6300 of consensus"
misc_feature    complement(31361. .31864) /note="L1MB6 repeat: matches 5938. .6108 of consensus"
repeat_region   32106. .32401          /note="match: GSS: Em:AQ120459"
repeat_region   33925. .34161          /note="L1MA2 repeat: matches 5747. .6050 of consensus"
repeat_region   34211. .34306          /note="MIR repeat: matches 13. .262 of consensus"
repeat_region   35045. .35108          /note="MIR repeat: matches 48. .146 of consensus"
misc_feature    35108                  /note="57 copies 2 mer at 76% conserved"
repeat_region   35119. .35267          /note="AluY repeat: matches 170. .311 of consensus"
repeat_region   35268. .35513          /note="L1ME3 repeat: matches 5891. .5951 of consensus"
repeat_region   35593. .35780          /note="AluY repeat: matches 13. .254 of consensus"
repeat_region   3652. .37065           /note="L1ME3 repeat: matches 5620. .5808 of consensus"
misc_feature    3692. .44026           /note="match: GSS: Em:AQ334107"
repeat_region   40160. .40610          /note="L1ME3 repeat: matches 5509. .6157 of consensus"
misc_feature    40377. .41698          /note="L2 repeat: matches 2629. .2709 of consensus"
repeat_region   44378. .44457          /note="L2 repeat: matches 2005. .2286 of consensus"
repeat_region   44718. .44985          /note="L1PA4 repeat: matches 3312. .4636 of consensus"
repeat_region   41699. .43199          /note="L1PA4 repeat: matches 4645. .6146 of consensus"
misc_feature    complement(47915. .48258) /note="match: GSS: Em:AQ059083"
repeat_region   43305. .43385          /note="AluSx repeat: matches 5. .305 of consensus"
repeat_region   48798. .49089          /note="104 copies 2 mer aa 76% conserved"
repeat_region   49133. .49340          /note="L1PA16 repeat: matches 5287. .5375 of consensus"
repeat_region   43386. .44026          /note="L1MC4 repeat: matches 6477. .6655 of consensus"
repeat_region   49343. .49521          /note="L1MC4 repeat: matches 1. .501 of consensus"
repeat_region   49522. .49998          /note="MLT2CB repeat: matches 6655. .7093 of consensus"
repeat_region   49999. .50436          /note="L1MC4 repeat: matches 6655. .7093 of consensus"
```

```
repeat_region   50869. .50974          /note="L1MC4 repeat: matches 7093. .7644 of consensus"
repeat_region   51005. .51316          /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region   51887. .52533          /note="L2 repeat: matches 1434. .2116 of consensus"
repeat_region   52544. .52819          /note="MLT2A repeat: matches 174. .453 of consensus"
repeat_region   53889. .54061          /note="MER96 repeat: matches 2. .173 of consensus"
repeat_region   55455. .55553          /note="L2 repeat: matches 2579. .2678 of consensus"
repeat_region   55646. .55719          /note="THE1C repeat: matches 291. .369 of consensus"
repeat_region   55720. .56032          /note="AluSx repeat: matches 1. .306 of consensus"
repeat_region   56033. .56348          /note="THE1C repeat: matches 7. .291 of consensus"
repeat_region   56955. .57266          /note="AluSx repeat: matches 1. .302 of consensus"
misc_feature    complement(57062. .57330) /note="match: GSS: Em:AQ487215"
```

```
Query Match     9.9%;  Score 86; DB 9; Length 57447;
Best Local Similarity 100.0%;  Pred. No. 7.5e-39;
  Matches  86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      39 GTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATCCGGCCGCCTCGG 98
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35420 GTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATCCGGCCGCCTCGG 35479

Qy      99 CCTCCCAAAGTGCTGGGATTACAGGC 124
           ||||||||||||||||||||||||||
Db   35480 CCTCCCAAAGTGCTGGGATTACAGGC 35505
```

Search completed: May 24, 2004, 17:39:20
Job time : 2505 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 15:49:17 ; Search time 333 Seconds
         (without alignments)
         11060.624 Million cell updates/sec

Title: US-10-621-363-12
Perfect score: 867
Sequence: 1 tcgagctttttttttttttt...........accaatctcgtgccgaattc 867
Scoring table: OLIGO_NUC
      Gapop 60.0 , Gapext 60.0
Searched: 3373863 seqs, 2124099041 residues
Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
   1: geneseqn1980s:*
   2: geneseqn1990s:*
   3: geneseqn2000s:*
   4: geneseqn2001as:*
   5: geneseqn2001bs:*
   6: geneseqn2002s:*
   7: geneseqn2003as:*
   8: geneseqn2003bs:*
   9: geneseqn2003cs:*
   10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	867	100.0	867	2	AAX30367	Aax30367 DNA encod
2	867	100.0	867	9	ADB47729	Adb47729 Novel hum
3	483	55.7	58181	7	ABZ68140	Abz68140 Human sec
4	483	55.7	58181	7	ABZ74619	Abz74619 Human sec
5	483	55.7	58181	9	ADC21010	Adc21010 Human sec
6	98	11.3	145831	6	ABL69213	Abl69213 Prostate
7	98	11.3	145831	6	ABL66806	Abl66806 Lung canc
8	98	11.3	145831	6	ABL68588	Abl68588 Kidney ca
9	98	11.3	145831	6	ABL62309	Abl62309 Colon ade
10	98	11.3	145831	6	ABT10149	Abt10149 Human bre
11	86	9.9	49561	4	AAK82012	Aak82012 Human imm
12	75	8.7	267	4	AAI62868	Aai62868 Human gen
13	74	8.5	579	4	AAH09234	Aah09234 Human CDN
14	74	8.5	1555	4	AAH15304	Aah15304 Human CDN
15	74	8.5	139904	6	ABK83562	Abk83562 Human ova
16	71	8.2	269	7	ABL87463	Abl87463 Human ova
17	71	8.2	201239	7	ACA64924	Aca64924 Human PlZ
18	69	8.0	634	4	AAK80474	Aak80474 Human imm
19	69	7.8	573	4	AAK80472	Aak80472 Human imm
20	68	7.8	635	4	AAK80470	Aak80470 Human imm
21	66	7.6	538	4	AAI86635	Aai86635 Human pol
22	66	7.6	700	4	AAH92579	Aah92579 Human inf
23	66	7.6	26657	6	AAI72317	Aai72317 Human tra
24	66	7.6	46366	4	AAK82098	Aak82098 Human imm
25	66	7.6	50000	9	ADC60735	Adc60735 Human sli
26	66	7.6	148567	7	ABS55500	Abs55500 Gene enco
27	66	7.6	148567	7	ACA62841	Aca62841 Human kin
28	65	7.5	1857	4	AAS42189	Aas42189 Genomic s
29	65	7.5	1857	4	AAK69783	Aak69783 Human lun
30	65	7.5	8663	4	AAS30133	Aas30133 Human lun
31	65	7.5	8663	9	ADB33470	Adb33470 Human nov
32	65	7.5	8663	9	ABK83568	Abk83568 Human DNA
33	65	7.5	201143	3	AAC18048	Aac18048 Human sec
34	64	7.4	192	3	AAL05992	Aal05992 Human rep
35	64	7.4	264	4	ABL98556	Abl98556 Human tes
36	64	7.4	264	4	AAK76486	Aak76486 Human imm
37	64	7.4	289	4	ABK42566	Abk42566 Genomic s
38	64	7.4	290	4	ABK42567	Abk42567 Genomic s
39	64	7.4	290	8	ADB60723	Adb60723 Connectiv
40	64	7.4	290	8	ADB60722	Adb60722 Connectiv
41	64	7.4	291	4	ABK42569	Abk42569 Genomic s
42	64	7.4	291	4	ABK42568	Abk42568 Genomic s
43	64	7.4	291	8	ADB60725	Adb60725 Connectiv
44	64	7.4	291	8	ADB60724	Adb60724 Connectiv
45	64	7.4	295	4	AAK81291	Aak81291 Human imm

ALIGNMENTS

RESULT 1
AAX30367
ID AAX30367 standard; DNA; 867 BP.
XX
AC AAX30367;
XX
DT 14-MAY-1999 (first entry)
XX
DE DNA encoding a human secreted protein.
XX
KW Secreted protein; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; foetal deficiency; blood disorder;
KW CNS disorder; immune system disease; autoimmune disease; hepatic disease;
KW renal disease; diabetes; inflammation; allergy; ischemic shock;
KW Alzheimer's; cognitive disorder; schizophrenia; cardiovascular disorder;
KW prostate disease; asthma; osteoporosis; arthritis; ss.
OS Homo sapiens.
XX
PN WO9907891-A1.
XX
PD 18-FEB-1999.
XX
PF 04-AUG-1998; 98WO-US016235.
XX
PR 05-AUG-1997; 97US-0054798P.
PR 05-AUG-1997; 97US-0054803P.
PR 05-AUG-1997; 97US-0054804P.
PR 05-AUG-1997; 97US-0054806P.
PR 05-AUG-1997; 97US-0054807P.
PR 05-AUG-1997; 97US-0054808P.
PR 05-AUG-1997; 97US-0054809P.
PR 05-AUG-1997; 97US-0053099P.
PR 05-AUG-1997; 97US-0053310P.
PR 05-AUG-1997; 97US-0053311P.
PR 05-AUG-1997; 97US-0053312P.
PR 05-AUG-1997; 97US-0055386P.
PR 18-AUG-1997; 97US-0055970P.
PR 19-AUG-1997; 97US-0055986P.
PR 19-AUG-1997; 97US-0056364P.
PR 19-AUG-1997; 97US-0056365P.
PR 19-AUG-1997; 97US-0056366P.
PR 19-AUG-1997; 97US-0056367P.
PR 19-AUG-1997; 97US-0056370P.
PR 19-AUG-1997; 97US-0056371P.
PR 19-AUG-1997; 97US-0056557P.

```
PR   19-AUG-1997;   97US-0055563P.
PR   19-AUG-1997;   97US-0056731P.
PR   19-AUG-1997;   97US-0056732P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM, Soppet DR, Ebner R, Olsen HS, Young PE, Greene JM;
PI   Ferrie AM, Yu G, Ni J, Rosen CA, Brewer LA, Janat F;
XX
DR   WPI; 1999-167452/14.
DR   P-PSDB; AAY10846.
XX
SQ
PS   Claim 3; Page 253-254; 331pp; English.
XX
XX   New isolated human genes encoding secreted polypeptides - useful for
PT   diagnosis and treatment of pathological diseases.
XX
CC   The specification describes secreted proteins and their corresponding
CC   polynucleotides which are useful for preventing, treating or ameliorating
CC   medical conditions, e.g by protein or gene therapy. Pathological
CC   conditions can also be diagnosed by determining the amount of the
CC   secreted polypeptides in a sample or by determining the presence of
CC   mutations in the polynucleotides. Specific uses are described for each of
CC   the products, based on which tissues they are most highly expressed in,
CC   and include developing products for the diagnosis or treatment of cancer,
CC   tumours, neurodegenerative disorders, developmental abnormalities and
CC   foetal deficiencies, blood disorders, CNS disorders, diseases of the
CC   immune system, autoimmune diseases, hepatic and renal disease, diabetes,
CC   inflammation, allergies, ischemic shock, Alzheimer's and cognitive
CC   disorders, schizophrenia, cardiovascular disorders, prostate diseases,
CC   asthma, diseases involving osteoclasts such as osteoporosis, arthritis
CC   or malignancies, diseases of testes, lung or thymus, digestive/endocrine
CC   disorders, infections and AIDS. The polypeptides are also useful for
CC   identifying their binding partners
XX
SQ   Sequence 867 BP; 219 A; 231 C; 222 G; 192 T; 0 U; 3 Other;

     Query Match   100.0%;  Score 867;  DB 2;  Length 867;
     Best Local Similarity   100.0%;   Pred. No. 5.4e-308;
     Matches  867;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       1 TCGAGTTTTTTTTTTTTTTTTTTTTTTAAGTAGAATGGGGTTTCACCGTGTTAGCCAGGAT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TCGAGTTTTTTTTTTTTTTTTTTTTTTAAGTAGAATGGGGTTTCACCGTGTTAGCCAGGAT  60

Qy      61 GGTCTCGGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTAC  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGTCTCGGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTAC  120

Qy     121 AGGCATGAGCCACTCCGCCCACGGCCGGTCTTTTTAAACATTCCCCAGGACTGGTACAGCCAA  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGGCATGAGCCACTCCGCCCACGGCCGGTCTTTTTAAACATTCCCCAGGACTGGTACAGCCAA  180

Qy     181 CCCATACTCACCTGACATTTGGGAACTCCCCCACGGCCCATAACTGAICTGCAGAGGTA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCCATACTCACCTGACATTTGGGAACTCCCCCACGGCCCATAACTGAICTGCAGAGGTA  240

Qy     241 AGACCAAGACAAGRAATGGGGGATTCACATCTAAGGTCTGGTGATGGCTGATGAAGGAAG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGACCAAGACAAGRAATGGGGGATTCACATCTAAGGTCTGGTGATGGCTGATGAAGGAAG  300

Qy     301 AAGAATCAGCGAACAAAAAGCCTCTAGGTCTTTCTTACCACAACACCTCTCTGCCCACCT  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAGAATCAGCGAACAAAAAGCCTCTAGGTCTTTCTTACCACAACACCTCTCTGCCCACCT  360

Qy     361 GCTTTGAAAGGGCGCAGAAGTATAGTGGGCAGCTGCCAGTGAAGGGATCT  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCTTTGAAAGGGCGCAGAAGTATAGTGGGCAGCTGCCAGTGAAGGGATCT  420

Qy     421 GGAGAAATACTCACACTTTGAGGGTGCTCGCCCCTCTTCATCAGCCAGCTCTAACTTAAGCC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAGAAATACTCACACTTTGAGGGTGCTCGCCCCTCTTCATCAGCCAGCTCTAACTTAAGCC  480

Qy     481 AATGACCCCACACGGGAGGCTTACACAAGTYCAAAACAGGGCCCAAAATGCATTCATTGAGCAGGGG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AATGACCCCACACGGGAGGCTTACACAAGTYCAAAACAGGGCCCAAAATGCATTCATTGAGCAGGGG  540

Qy     541 GAGGCCAAGGACTCCGGAGGAGGAGAGGGCCGGAATTAAGGCTGGGTGGTCTATTTCCGATCCATA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAGGCCAAGGACTCCGGAGGAGGAGAGGGCCGGAATTAAGGCTGGGTGGTCTATTTCCGATCCATA  600

Qy     601 GAGAGAGGAGGTGGGCAGGGCCCTTTGATTAATGTRATCATTCTTGAATGCAAGCTTCA  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GAGAGAGGAGGTGGGCAGGGCCCTTTGATTAATGTRATCATTCTTGAATGCAAGCTTCA  660

Qy     661 AAATCCGGGTTATTCCCGGGGTGAGAATTGGACGGACTAACCACCTGGGTGTCATGGCCAAGCCT  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AAATCCGGGTTATTCCCGGGGTGAGAATTGGACGGACTAACCACCTGGGTGTCATGGCCAAGCCT  720

Qy     721 CCAGGGCCCGACTGGCCCAGAGACAGATCCGCAAGAGGCTTCTGCAGCCCAGGCTTCTGGTTCGGTGCCAA  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CCAGGGCCCGACTGGCCCAGAGACAGATCCGCAAGAGGCTTCTGCAGCCCAGGCTTCTGGTTCGGTGCCAA  780

Qy     781 GCCACTCGGATTTGAACCCCGGCGTCCTCAAAGGTCAGCTGTGTAGCCTTGAACTGAATGAAYCACC  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GCCACTCGGATTTGAACCCCGGCGTCCTCAAAGGTCAGCTGTGTAGCCTTGAACTGAATGAAYCACC  840

Qy     841 TGCTATGACCAAATCTCGTGCCCGGAATTC  867
           ||||||||||||||||||||||||||||||
Db     841 TGCTTATGACCAAATCTCGTGCCCGGAATTC  867
```

```
RESULT 2
ADB47729
ID   ADB47729 standard; cDNA; 867 BP.
XX
AC   ADB47729;
XX
XX
DT   04-DEC-2003  (first entry)
XX
DE   Novel human secreted protein cDNA #2.
XX
KW   ss; gene; human; secreted protein; insulin; haemoglobin S; haemoglobin B;
KW   superoxidase; SOD; catalase; DNA repair protein; oncogene;
KW   tumour suppressor; tumour necrosis factor; TNF; inflammation;
KW   blood vessel growth inhibition; neoplasm; immune response; immune system disorder;
KW   hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW   peripheral artery disease; limb ischaemia; arterio-arterial fistula;
KW   arteriovenous fistula; congenital heart defect;
KW   neovascularisation disorder; wound healing;
KW   epithelial cell proliferation; neurological disease; Alzheimer's disease;
KW   Parkinson's disease; Huntington's disease; mania; dementia;
KW   infectious disease.
XX
OS   Homo sapiens.
XX
PN   US2003054443-A1.
XX
PD   20-MAR-2003.
XX
PF   04-OCT-2001;  2001US-00969730.
XX
PR   05-AUG-1997;  97US-0054798P.
PR   05-AUG-1997;  97US-0054803P.
PR   05-AUG-1997;  97US-0054804P.
PR   05-AUG-1997;  97US-0054806P.
PR   05-AUG-1997;  97US-0054807P.
PR   05-AUG-1997;  97US-0054808P.
PR   05-AUG-1997;  97US-0054809P.
PR   05-AUG-1997;  97US-0055309P.
PR   05-AUG-1997;  97US-0055310P.
PR   05-AUG-1997;  97US-0055311P.
PR   05-AUG-1997;  97US-0055312P.
PR   05-AUG-1997;  97US-0055386P.
PR   18-AUG-1997;  97US-0055970P.
PR   18-AUG-1997;  97US-0055986P.
```

PR   19-AUG-1997; 97US-0056364P.
PR   19-AUG-1997; 97US-0056365P.
PR   19-AUG-1997; 97US-0056366P.
PR   19-AUG-1997; 97US-0056367P.
PR   19-AUG-1997; 97US-0056370P.
PR   19-AUG-1997; 97US-0056371P.
PR   19-AUG-1997; 97US-0055557P.
PR   19-AUG-1997; 97US-0055563P.
PR   19-AUG-1997; 97US-0055731P.
PR   19-AUG-1997; 97US-0057731P.
PR   19-AUG-1997; 97US-0055732P.
PR   04-AUG-1998; 98WO-US016235.
PR   04-FEB-1999; 99US-0024411Z.
PR   06-OCT-2000; 2000US-0238291P.
PR   01-FEB-2001; 2001US-0077463S.
XX
PA   (RUBE/) RUBEN S M.
PA   (SOPP/) SOPPET D R.
PA   (EBNE/) EBNER R.
PA   (OLSE/) OLSEN H S.
PA   (YOUN/) YOUNG P E.
PA   (GREE/) GREENE J M.
PA   (FERR/) FERRIE A M.
PA   (YUGG/) YU G.
PA   (NIJ/)   NI J.
PA   (ROSE/) ROSEN C A.
PA   (BREW/) BREWER L A.
PA   (JANA/) JANAT F.
PA   (BIRS/) BIRSE C E.
XX
PI   Ruben SM, Soppet DR, Ebner R, Olsen HS, Young PE, Greene JM,
PI   Ferrie AM, Yu G, Ni J, Rosen CA, Brewer LA, Janat F, Birse CE;
XX
DR   WPI; 2003-695903/66.
DR   P-PSDB; ADB47826.
XX
PN   Novel human secreted proteins useful for treating and/or diagnosing
PT   disorders of immune system, cardiovascular disorders such as peripheral
PT   artery disease, neurological diseases such as Alzheimer's disease.
XX
PS   Claim 3; Page 205; 333pp; English.

CC   The invention relates to novel human secreted proteins. The protein is
CC   useful for preventing, treating or ameliorating a medical condition. The
CC   protein is useful for diagnosing a pathological condition or
CC   susceptibility to pathological condition in a subject. The protein is
CC   useful for identifying a binding partner. The nucleic acid is useful for
CC   diagnosing pathological condition or a susceptibility to pathological
CC   condition in a subject. The protein is useful as reagents for
CC   differential identification of the tissues or cell types present in a
CC   biological sample. The protein can be administered to patients having
CC   absent or decreased levels of polypeptides e.g. insulin, to supplement
CC   absent or decreased levels of different polypeptides e.g. haemoglobin s
CC   for haemoglobin B, superoxidase (SOD), catalase, DNA repair protein, to
CC   inhibit the activity of a polypeptide e.g. an oncogene or tumour
CC   suppressor, to reduce the activity of polypeptide e.g. by binding to a
CC   receptor, to activate the activity of membrane bound receptor by competing
CC   with it for free ligand e.g. soluble tumour necrosis factor (TNF)
CC   receptors used in reducing inflammation, or to bring about a desired
CC   response e.g. blood vessel growth inhibition, enhancement of immune
CC   response to proliferative cells or tissues. The protein and the nucleic
CC   acid are useful for treating, preventing, detecting, diagnosing disorders
CC   of immune system involving abnormal growth of specific types of cells as
CC   well as of other cell types where expression has been observed. The
CC   protein, the nucleic acid and antibodies are useful for treating,
CC   preventing and/or diagnosing diseases, disorders and/or conditions of
CC   immune system, hyperproliferative disorders including neoplasms,
CC   cardiovascular disorders (such as peripheral artery disease, limb
CC   ischaemia, arterio-arterial fistula, arteriovenous fistula, congenital
CC   heart defects, etc), neovascularisation disorders, wound healing and
CC   epithelial cell proliferation, neurological diseases (such as Alzheimer's
CC   disease, Parkinson's disease, Huntington's disease, mania, dementia,
CC   etc), infectious diseases caused by virus, bacteria, fungi, etc. The
CC   present sequence represents a cDNA encoding a novel human secreted

CC   protein.
XX
SQ   Sequence 867 BP; 219 A; 231 C; 222 G; 192 T; 0 U; 3 Other;

```
Qy          Query Match              100.0%;   Score 867;  DB 9;  Length 867;
            Best Local Similarity    100.0%;   Pred. No. 5.4e-308;
Matches   867;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCGGAGTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATAGGGGTTTCACCGTGTTAGCCAGGAT         60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGGAGTTTTTTTTTTTTTTTTTTTAAGTAGAGAGATAGGGGTTTCACCGTGTTAGCCAGGAT         60

Qy         61 GGTCTCGGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTAC        120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGTCTCGGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTAC        120

Qy        121 AGGCATGAGCCACTGCGCCCCAGCCGGGTGCGCGTGCTTTTTAAACATTCCCAGGAGCCAA        180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCATGAGCCACTGCGCCCCAGCCGGGTGCGCGTGCTTTTTAAACATTCCCAGGAGCCAA        180

Qy        181 CCCATACTCACCTGACATTTGAGAGAACTCCCCCCCCCACCGGCCCATAACTGATCTGCAGAGGTA        240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCATACTCACCTGACATTTGAGAGAACTCCCCCCCCCACCGGCCCATAACTGATCTGCAGAGGTA        240

Qy        241 AGACCAAGACAAGAAATGGGGATTCACATCTAAGGTCTGGTGATGGCTGATGAGGAAG        300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGACCAAGACAAGAAATGGGGATTCACATCTAAGGTCTGGTGATGGCTGATGAGGAAG        300

Qy        301 AAGAATCAGCGAACAAAAGCCTCTAGGTGTCTTTTCTTACCACAACAAACACTTCTTCTGCCCACCT        360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGAATCAGCGAACAAAAGCCTCTAGGTGTCTTTTCTTACCACAACAAACACTTCTTCTGCCCACCT        360

Qy        361 GCTTTGAAAGGGGCGCAGAAGTATAGTAGTGGGCGGAGCTGCCACCTGCTACAGTGAAGGGATCT        420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTTGAAAGGGGCGCAGAAGTATAGTAGTGGGCGGAGCTGCCACCTGCTACAGTGAAGGGATCT        420

Qy        421 GGAGAAAATACTCACACTTTGAGGTGTCGCCCTCTTCATCAGCCAGCCTTAACTTAAGCC        480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAGAAAATACTCACACTTTGAGGTGTCGCCCTCTTCATCAGCCAGCCTTAACTTAAGCC        480

Qy        481 AATGACCCCACGGGAGCTTACACAAGTYCAAACAGGGCCCAAAATCATTCATGAGCAGGGG        540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATGACCCCACGGGAGCTTACACAAGTYCAAACAGGGCCCAAAATCATTCATGAGCAGGGG        540

Qy        541 GAGGCCAAAGGACTCCGGAGGAGGAGGCCCAATAAGGCTGGTGGTGCTATTTCCGATCATA        600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGCCAAAGGACTCCGGAGGAGGAGGCCCAATAAGGCTGGTGGTGCTATTTCCGATCATA        600

Qy        601 GAGAGCAGAGGTGGGCAGGCCCTTTGATTAAGTATCATTCTTGAATGCAAGCTTCA        660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAGCAGAGGTGGGCAGGCCCTTTGATTAAGTATCATTCTTGAATGCAAGCTTCA        660

Qy        661 AAATCCGGGTATGCCGGGTGAGAATGAGCAGGAGTAACACCTGGGTGTCATGGCCAAGCCT        720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAATCCGGGTATGCCGGGTGAGAATGAGCAGGAGTAACACCTGGGTGTCATGGCCAAGCCT        720

Qy        721 CCAGGGCCCGACTGGCCAGAGACAGATCCGCAACCGGGCTCTGGTGCCAA        780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCAGGGCCCGACTGGCCAGAGACAGATCCGCAACCGGGCTCTGGTGCCAA        780

Qy        781 GCCACTCGGATTTGAAACCCGGGCTTCCTCAAGGTCAGCTGTGTAGCCCTTGANTGAAYCACC        840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCACTCGGATTTGAAACCCGGGCTTCCTCAAGGTCAGCTGTGTAGCCCTTGANTGAAYCACC        840

Qy        841 TGCTATGACCAATCTGTGCCGGAATTC   867
              |||||||||||||||||||||||||||
Db        841 TGCTATGACCAATCTGTGCCGGAATTC   867
```

RESULT 3
AB268140/c
ID   AB268140 standard; DNA; 58181 BP.
XX

```
AC   AB268140;
XX
DT   26-MAR-2003 (first entry)
XX
DE   Human secreted protein encoding genomic DNA SEQ ID NO 1663.
XX
KW   Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW   vinculde; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW   vulnerary; antibacterial; antiparkinsonian; antisickling; antianaemic;
KW   antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW   antiinflammatory; antiallergic; antidiabetic; antiulcer; anticonvulsant;
KW   antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW   cardiovascular disorder; neurological disease; nephrotropic;
KW   gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO20277186-A2.
XX
PD   03-OCT-2002.
XX
PF   26-MAR-2002; 2002WO-US009188.
XX
PR   27-MAR-2001; 2001US-0278650P.
PR   12-SEP-2001; 2001US-00950082.
PR   12-SEP-2001; 2001US-00950083.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2003-040583/03.
XX
PT   New human secreted proteins encoded by genes contained in cDNA clones
PT   (e.g. HGCAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT   multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT   West Nile fever.
XX
PS   Disclosure; Page 2272-2286; 2423pp; English.
XX
CC   The invention relates to novel human genes (AB268891-AB268209) and the
CC   encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC   treating or ameliorating medical conditions e.g. by protein or gene
CC   therapy. The genes are isolated from a range of human tissues disclosed
CC   in the specification. The nucleic acids, proteins, antibodies and
CC   (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC   (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC   adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC   lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC   myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC   cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC   bacterial, fungal and parasitic infections
XX
SQ   Sequence 58181 BP; 15503 A; 12755 C; 13504 G; 16419 T; 0 U; 0 Other;

Query Match             55.7%;  Score 483;  DB 7;  Length 58181;
Best Local Similarity   99.5%;  Pred. No. 7.9e-168;
Matches   823;  Conservative   0;  Mismatches   2;  Indels   2;  Gaps   2;

Qy          6 TTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTTCACGGTGTTAGCCAGGATGGTCT 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52565 TTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGGTTTCACGGTGTTAGCCAGGATGGTCT 52506

Qy         66 CGAACTCTTGACCTCGTGATCCCGCCGCCTCCCAAAGTGCTGGGATTACAGGCA 125
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52505 CGAACTCCTGACCTCGTGATCCCGCCGCCTCCCAAAGTGCTGGGATTACAGGCA 52446

Qy        126 TGAGCCACTGCGCGCCCAGCGCGGGTCTTTTTAAACAATTCCCCAGGACACTGTACAGCCAACCCAT 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52445 TGAGCCACTGCGCGCCCAGCGCGGTCTTTTTAAACAATTCCCCAGGACACTGTACAGCCAACCCAT 52386
```

```
Qy        186 ACTCACCTGACACATTTGGGAGAACTCCCCCCCACGGCGGCCCATAACTGATCTGCAGAGGTAAGACC 245
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52385 ACTCACCTGACACATTTGGGAGAACTCCCCCCCACGGCGGCCCATAACTGATCTGCAGAGGTAAGACC 52326

Qy        246 AAGAGGCAAGAAATGGGGGGATTCACATCTAAGGTCTGGTGATGGGCTGATGAAGGAGGAAGAA 305
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52325 AAGAGCAAGAAATGGGGGGATTCACATCTAAGGTCTGGTGATGGGCTGATGAAGGAGGAAGAA 52266

Qy        306 TCAGCGGACAAAAAGCCTCTTAGGTCTTTTCTTACCACAACACCTTCTTCGCCCACCTGCTTT 365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52265 TCAGCGGACAAAAAGCCTCTTAGGTCTTTTCTTACCACAACACCTTCTTCGCCCACCTGCTTT 52206

Qy        366 GAAAGGCGGCAGGAGTAGTAGGGCGGA-GCTGCCACCTCGTACAGTGAAGGGGATCTGGAG 424
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db      52205 GAAAGGCGGCAGGAGTAGTAGGGCGGAGGCTGCCACCTCGTACAGTGAAGGGGATCTGGAG 52146

Qy        425 AAAATACTCACCACTTTGAGGTGCTCGCCCTTCATCAGCCAGCTCTAACTTAAGCCAATG 484
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52145 AAAATACTCACCACTTTGAGGTGCTCGCCCTTCATCAGCCAGCTCTAACTTAAGCCAATG 52086

Qy        485 ACCCCACGGGGAG-CTTACACAAGTYCAAACAGGGCCCAAAATGCATTCATGAGCAGGGGGGAG 543
              ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      52085 ACCCCACGGGGAGCCTTACACAAGTCCAAACAGGGCCCAAAATGCATTCATGAGCAGGGGGGAG 52026

Qy        544 GCCAAAGGAGACTCCGGAGGAGGAGGGCCCAATAAGGCTGGTGCTATTTCCGATCCATCATAGAG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52025 GCCAAAGGAGACTCCGGAGGAGGAGGGCCCAATAAGGCTGGTGCTATTTCCGATCCATCATAGAG 51966

Qy        604 AGAGCAGGAGGTGGGGCAGGCCCCTTTTGATTAAATGTATCATTCTTTGAAATGCAAGGTTCAAA 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51965 AGAGCAGGAGGTGGGGCAGGCCCCTTTTGATTAAATGTATCATTCTTTGAAATGCAAGGTTCAAA 51906

Qy        724 GGGGCGGACTGGGCCCAGAGACAGATCCGCAAGAGAGGCTCTCGAGCCAGCGTTGGGTTGCCAAGCC 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51845 GGGGCGGACTGGGCCCAGAGACAGATCCGCAAGAGAGGCTCTCGAGCCAGCGTTGGGTTGCCAAGCC 51786

Qy        784 ACTGGGATTGAACCCCGGGCTCCTCAAGGTCAGCTGTGTAGCCTTGA 830
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      51905 ACTGGGATTGAACCCCGGGCTCCTCAAGGTCAGCTGTGTAGCCTTGA 51846
```
(note: some alignment rows above may be read imperfectly)

```
AC   AB274619;
XX
ID   AB274619 standard; DNA; 58181 BP.
XX
AC   AB274619;
XX
DT   12-MAY-2003 (first entry)
XX
DE   Secreted protein gene 372 genomic fragment HWBCN36, SEQ ID NO:1766.
XX
KW   Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW   autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW   acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW   drug screening; chromosome identification; chromosome mapping;
KW   cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW   antianaemic; vulnerary; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO20277013-A2.
XX
PD   03-OCT-2002.
XX
PF   26-MAR-2002; 2002WO-US009370.
XX
PR   27-MAR-2001; 2001US-0278650P.
PR   12-SEP-2001; 2001US-00950082.
```

RESULT 4

PR   12-SEP-2001; 2001US-00950083.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2003-040578/03.
XX
PT   New human secreted proteins and nucleic acids, useful for detecting or
PT   treating cancer or other hyperproliferative disorders, autoimmune
PT   disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
XX
PS   Disclosure; Page 2331-2345; 2474pp; English.
XX
CC   AB273281-AB273697 represent cDNAs corresponding to 391 human secreted
CC   protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC   AB273698-AB274687 represent human secreted protein genomic fragments. The
CC   invention also encompasses antibodies specific for the secreted proteins,
CC   the use of the secreted proteins in drug screening and recombinant
CC   identification, chromosome mapping, in gene therapy, for identifying
CC   individuals from minute biological samples, as hybridisation probes, and
CC   as molecular weight markers. The present sequence represents a human
CC   secreted protein genomic fragment referred to in the disclosure of the
CC   invention.
CC   vectors and host cells comprising a nucleic acid of the invention. The
CC   secreted proteins are thought to be involved in biological activities
CC   associated with cellular signaling, cellular differentiation, cell
CC   migration, prohormone activation and neurotransmitter activity. The
CC   secreted proteins, nucleic acids encoding them, antibodies or antibody
CC   fragments specific for the secreted proteins, and modulators of protein
CC   activity are useful for diagnosing or treating cancers or other
CC   hyperproliferative disorders. Additionally, the secreted proteins and
CC   their nucleic acids may also be used in the treatment of autoimmune
CC   disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC   (acquired immunodeficiency syndrome); hepatitis, anaemia, and to promote
CC   wound healing. Nucleic acids of the invention may be used for chromosome
XX
SQ   Sequence 58181 BP; 15503 A; 12755 C; 13504 G; 16419 T; 0 U; 0 Other;

Query Match   55.7%; Score 483; DB 7; Length 58181;
Best Local Similarity   99.5%; Pred. No. 7.9e-168;
Matches 823; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```
Qy          6 TTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGGATGAGAGGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52565 TTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAGGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 52506

Qy         66 CGAACTCCTTGACCTCGTGATCCGCCCGCGCCCGCCTCGGCGTTCCCAAAGTTGGGATTACAGGCA 125
              ||||||||| ||||||||||||||| |||||||||||| |||||||||||||||||||||||||
Db      52505 CGAACTCCTCGACCTCGTGATCCGCCTCCCGGCCTCCCAAAGTGCCTGGGATTACAGGCA 52446

Qy        126 TGAGCCACTGCGCCCCAGCCCGGTCTTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52445 TGAGCCACTGCGCCCCAGCCGGTCTTCTTTTTAAACATTCCCCAGGACTGTACAGCCAACCCAT 52386

Qy        186 ACTCACCTTGACATTTGGGGAACTCCCCCCCCACGCGCCATAACTGATCTGCGAGAGGTAAGACC 245
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52385 ACTCACCTGACATTTGGGGAACTCCCCCCCCACACGGCCATAACTGATCTGCGAGAGGTAAGACC 52326

Qy        246 AAGAGCAAGAAATGGGGGGATTCACATCAATCTAAGGTCTCGGTGATGGCTCGAATGAAGGAAGAAGGA 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52325 AAGAGCAAGAAATGGGGGGATTCAATCTAAGGTCTCGGTGATGGCTCGAATGAAGGAAGGAAGAAGAA 52266

Qy        306 TCAGCGGAACAAAAGCCTCTAGGTCTTTCTTACCACAACACCTTCTCTGCCCACCTGCTTT 365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52265 TCAGCGGAACAAAAGCCTCTAGGTCTTTCTTACCACAACACCTTCTCTGCCCACCTGCTTT 52206

Qy        366 GAAAGGGGGCAGAAGTATAGTGGGGGGA-GCTGCCCACCTCTACAGTGAAGGGGATCTGGAG 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52205 GAAAGGGGGCAGAAGTATAGTGGGGGGACTGCCCACCCTCTACAGTGAAGGGGATCTGGAG 52146

Qy        425 AAATACTCACACTTTGAGGTGGCTCGGCCCCTCTTCATCAGCCAGCTCTAACTTAAGCCAATG 484
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52145 AAATACTCACACTTTGAGGTGGCTCGGCCCCTCTTCATCAGCCAGCTCTAACTTAAGCCAATG 52086
```

```
Db      52145 AAATACTCACACTTTGAGGTGGCTCGGCCCCTCTTCATCAGCCAGCTCTAACTTAAGCCAATG 52086
Qy        485 ACCCCACGGCGAG-CTTACACAAGTYCAAACACAGGCCCCAAATGCATACATGAGCAGGGGAG 543
              |||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||
Db      52085 ACCCCACGGCGAGCCTTACACAAGTCCAAACACAGGCCCCAAATGCATACATGAGCAGGGGAG 52026

Qy        544 GCCAAAGGACTCCGGAGAGGAGAGGGCCCAATAAGGCTGGTGTCGTATTTCCGATTCATAGAG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52025 GCCAAAGGACTCCGGAGAGGAGAGGGCCCAATAAGGCTGGTGTCGTATTTCCGATTCATAGAG 51966

Qy        604 AGAGCAGGAGGTGGGCAGGGCCCAGGGCCCTTTGATTAAATGTATCATTCTTGAAATGCAAAG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51965 AGAGCAGGAGGTGGGCAGGGCCCAGGGCCCTTTGATTAAATGTATCATTCTTGAAATGCAAAG 51906

Qy        664 TCCGGGTATCGCGGGTGAGGAATGAGCAGGACTAACACCTGGGTGTCATGGCAAGGCTCCA 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51905 TCCGGGTATCGCGGGTGAGGAATGAGCAGGACTAACACCTGGGTGTCATGGCAAGGCTCCA 51846

Qy        724 GGGCCGACTGGCCAGACAGATCCGCAAGAGGCTCTGCAGCCTCTGGTGCCAAGCC 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51845 GGGCCGACTGGCCAGACAGATCCGCAAGAGGCTCTGCAGCCTCTGGTGCCAAGCC 51786

Qy        784 ACTCGGATTTGAACCCCGGCCGCTCCTTCAAGGTCAGCTGTGTTGAGCCTTGA 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51785 ACTCGGATTTGAACCCCGGCCGCTCCTTCAAGGTCAGCTGTGTTGAGCCTTGA 51739
```

RESULT 5
ADC21010/c
ID   ADC21010 standard; DNA; 58181 BP.
XX
AC   ADC21010;
XX
DT   18-DEC-2003 (first entry)
XX
DE   Human secreted protein-related DNA sequence #428.
XX
KW   gene therapy; human; secreted protein; haemopoietic disorder;
KW   haematological disorder; anaemia; haemophilia; inflammatory disorder;
KW   inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
KW   leukaemia; wound healing; epithelial cell proliferation disorder;
KW   immune disorder; autoimmune disorder; asthmatic disorder;
KW   cardiovascular disorder; atherosclerosis; myocarditis;
KW   infectious disease; HIV; AIDS; endocrine disorder; diabetes;
KW   gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.
XX
OS   Homo sapiens.
XX
XX
PN   WO200292787-A2.
XX
PD   21-NOV-2002.
XX
PF   26-MAR-2002; 2002WO-US009257.
XX
PR   27-MAR-2001; 2001US-02786502P.
PR   12-SEP-2001; 2001US-00950082.
PR   12-SEP-2001; 2001US-00950083.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2003-122287/12.
XX
PT   New human secreted proteins and nucleic acid molecules, useful for
PT   preparing a diagnostic or pharmaceutical composition for diagnosing,
PT   preventing or treating hematopoietic or hematologic disorders, e.g.
PT   anemia or hemophilia.
XX
PS   Disclosure; SEQ ID NO 964; 1512pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of human
CC   secreted proteins. The DNA and protein sequences of the invention are

```
CC   useful for detecting, preventing, diagnosing, prognosticating, treating
CC   or ameliorating: haematopoietic or haematological disorders (e.g. anaemia
CC   and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
CC   and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
CC   wound healing and disorders of epithelial cell proliferation; immune
CC   disorders (e.g. autoimmune disorders and asthmatic disorders);
CC   cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC   infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC   and gastrointestinal disorders (e.g. duodenal ulcers and
CC   gastroenteritis). The present DNA sequence was used in the
CC   exemplification of the invention.
XX
SQ   Sequence 58181 BP; 15503 A; 12755 C; 13504 G; 16419 T; 0 U; 0 Other;
```

Query Match   55.7%;   Score 483; DB 9; Length 58181;
Best Local Similarity   99.5%;   Pred. No. 7.9e-168;
Matches 823; Conservative   0; Mismatches   2; Indels   2; Gaps   2;

```
Qy        6 TTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 65
            ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    52565 TTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCT 52506

Qy       66 CGAACTCCTGACCTCGTGATCCCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 125
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52505 CGAACTCCTGACCTCGTGATCCCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCA 52446

Qy      126 TGAGCCACTGCGGCCCAGCGGCCAGCCGGTCTTTTTAAACATTCCCAGGACGTACAGCCAACCCAT 185
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52445 TGAGCCACTGCGGCCCAGCGGCCAGCCGGTCTTTTTAAACATTCCCAGGACGTACAGCCAACCCAT 52386

Qy      186 ACTCACCTGACATTTGGGAACTCCCCCCCCACGGCCCATAACTGATCTGCAGAGGTAAGACC 245
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52385 ACTCACCTGACATTTGGGAACTCCCCCCCCACGGCCCATAACTGATCTGCAGAGGTAAGACC 52326

Qy      246 AAGAGCAAGAATGGGGGATTCACATCTAAGGTGTCTTGGTGTGATGAAGGAAGAAGAA 305
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52325 AAGAGCAAGAATGGGGGATTCACATCTAAGGTGTCTTGGTGTGATGAAGGAAGAAGAA 52266

Qy      306 TCAGCGGACAAAAAGCCCTCTAGGTGTCTTTTCTTACCACAACACCTCTTCGCCCACCTGCTTT 365
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52265 TCAGCGGACAAAAAGCCCTCTAGGTGTCTTTTCTTACCACAACACCTCTTCGCCCACCTGCTTT 52206

Qy      366 GAAAGGGGCCAGAAGTATAGTGGGGGA-GCTGCCCACCTTGCTACAGTGCAGTGAAGGGAATCTCGAG 424
            |||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db    52205 GAAAGGGGCAGAAGTATAGTGGGGCCAGGACTCGCCCACCTTGCTACAGTGCAGTGAAGGGAATCTGGAG 52146

Qy      425 AAATACTCACACTTTGAGGTGCTCGCCCCTTCTTCATCAGCCAGCCTCAACTTAAGCCAATG 484
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52145 AAATACTCACACTTTGAGGTGCTCGCCCCTTCTTCATCAGCCAGCCTCAACTTAAGCCAATG 52086

Qy      485 ACCCCACGGGAG-CTTACACAGTYCAAACAGGCCCAAATGCATTCATTCTTGAGCAGGGGGAG 543
            |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    52085 ACCCCACGGGAGGCTTACACAGTCCAAACAGGCCCAAATGCATTCATTCTTGAGCAGGGGGAG 52026

Qy      544 GCCAAAAGGACTCCCGAGGAGAGGACTAACACCTGGGGTGTCATGGGCAAGCCTTCA 603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    52025 GCCAAAAGGACTCCCGAGGAGAGGACTAACACCTGGGGTGTCATGGGCAAGCCTTCA 51966

Qy      604 AGAGCAGAGAGGTGGGGCAGGCCCTTTGATTAATGTTATCATTCTTGAATGCAAGCTTCAAA 663
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    51965 AGAGCAGAGAGGTGGGGCAGGCCCTTTGATTAATGTTATCATTCTTGAATGCAAGCTTCAAA 51906

Qy      664 TCCGGGTATTGCCCGGGTGAAATGAGCAGAGGACTAACACCTGGGGTGTCATGGGCAAGCCCTCCA 723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    51905 TCCGGGTA-TGCCGGGTGAGAATGAGCAGAGGACTAACACCTGGGGTGTCATGGGCAAGCCCTCCA 51846

Qy      724 GGGCGGACTGGGCCAAGAGACAGGAATCCGCAAGAGGCTCTGGAGCCAGCCTCGGTCGCCAAGCC 783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    51845 GGGCGGACTGGGCCAAGAGACAGGAATCCGCAAGAGGCTCTGGAGCCAGCCTCGGTCGCCAAGCC 51786

Qy      784 ACTCGGATTTGAACCCCGGCGCTTCCTCAAGGTCAGCTGTGTTAGCCTTGA 830
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    51785 ACTCGGATTTGAACCCCGGCGCTTCCTCAAGGTCAGCTGTGTTAGCCTTGA 51739
```

DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 7550; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
CC
SQ   Sequence 145831 BP; 37379 A; 35440 C; 35158 G; 37854 T; 0 U; 0 Other;

Query Match   11.3%; Score 98; DB 6; Length 145831;
Best Local Similarity   100.0%; Pred. No. 5.1e-27;
Matches   98; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGGGGTTTCACCGGTGGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGGGGTTTCACCGGTGGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGCCTCGGCCCTCCCAAAGTGCTGGGATTACAGGC 124
           |||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCCTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 7
ABL66806/c
ID   ABL66806 standard; DNA; 145831 BP.
XX
AC   ABL66806;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Lung cancer related gene sequence SEQ ID NO:5143.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US010838.
XX
PR   05-JUN-2000; 2000US-0209473P.
PR   05-JUN-2000; 2000US-0209531P.
PR   18-SEP-2000; 2000US-0233133P.
PR   18-SEP-2000; 2000US-0233617P.
PR   20-SEP-2000; 2000US-0234009P.
PR   20-SEP-2000; 2000US-0234034P.
PR   20-SEP-2000; 2000US-0234052P.
PR   22-SEP-2000; 2000US-0234509P.
PR   22-SEP-2000; 2000US-0234567P.
PR   25-SEP-2000; 2000US-0234923P.
PR   25-SEP-2000; 2000US-0234924P.
PR   25-SEP-2000; 2000US-0235077P.
PR   25-SEP-2000; 2000US-0235082P.
PR   25-SEP-2000; 2000US-0235134P.
PR   25-SEP-2000; 2000US-0235280P.
PR   26-SEP-2000; 2000US-0235637P.
PR   26-SEP-2000; 2000US-0235638P.
PR   27-SEP-2000; 2000US-0235711P.
PR   27-SEP-2000; 2000US-0235720P.
PR   27-SEP-2000; 2000US-0235840P.
PR   27-SEP-2000; 2000US-0235863P.
PR   28-SEP-2000; 2000US-0236028P.
PR   28-SEP-2000; 2000US-0236032P.
PR   28-SEP-2000; 2000US-0236033P.
PR   28-SEP-2000; 2000US-0236034P.
PR   28-SEP-2000; 2000US-0236109P.
PR   28-SEP-2000; 2000US-0236111P.
PR   29-SEP-2000; 2000US-0236842P.
PR   29-SEP-2000; 2000US-0236891P.
PR   02-OCT-2000; 2000US-0371720P.
PR   02-OCT-2000; 2000US-0371728P.
PR   02-OCT-2000; 2000US-0371733P.
PR   02-OCT-2000; 2000US-0372172P.
PR   02-OCT-2000; 2000US-0372778P.
PR   02-OCT-2000; 2000US-0372794P.
PR   02-OCT-2000; 2000US-0373295P.
PR   02-OCT-2000; 2000US-0373316P.
PR   03-OCT-2000; 2000US-0374252P.
PR   03-OCT-2000; 2000US-0375988P.
PR   03-OCT-2000; 2000US-0376042P.
PR   03-OCT-2000; 2000US-0376606P.
PR   03-OCT-2000; 2000US-0376608P.
PR   01-NOV-2000; 2000US-0244867P.
PR   01-NOV-2000; 2000US-0245084P.
XX
PI   Young PB, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
PA   (AVAL-) AVALON PHARM.
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 5143; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
SQ   Sequence 145831 BP; 37379 A; 35440 C; 35158 G; 37854 T; 0 U; 0 Other;

Query Match   11.3%; Score 98; DB 6; Length 145831;
Best Local Similarity   100.0%; Pred. No. 5.1e-27;
Matches   98; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGGGGTTTCACCGGTGGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   38600 AGTAGAGGATGGGGTTTCACCGTGTGTTAGCCAGGATGGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGGCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGC 124
            ||||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGGCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 8

AC   ABL68588;
XX
XX
DT   15-MAY-2002 (first entry)
XX
DE   Kidney cancer related gene sequence SEQ ID NO:6925.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US010838.
XX

PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 6925; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. M1 has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour.
XX
SQ   Sequence 145831 BP; 37379 A; 35440 C; 35158 G; 37854 T; 0 U; 0 Other;

Query Match   11.3%; Score 98; DB 6; Length 145831;
Best Local Similarity 100.0%; Pred. No. 5.1e-27;
Matches   98; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGGATGGGGTTTCACCGTGTGTTAGCCAGGATGGGTCTCGAACTCCTGACCTCGTGATC 86
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGGATGGGGTTTCACCGTGTGTTAGCCAGGATGGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGGCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGC 124
            ||||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGGCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 9

AC   ABL62309;
XX
XX
DT   15-MAY-2002 (first entry)
XX
DE   Colon adenocarcinoma related gene sequence SEQ ID NO:646.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US010838.
XX

ID   ABL62309 standard; DNA; 145831 BP.
XX
AC   ABL62309;
XX
RESULT 9
ABL62309/c

PR   05-JUN-2000; 2000US-0209473P.
PR   05-JUN-2000; 2000US-0209531P.
PR   18-SEP-2000; 2000US-0231133P.
PR   18-SEP-2000; 2000US-0233617P.

PR 20-SEP-2000; 200OUS-0234009P.
PR 20-SEP-2000; 200OUS-0234034P.
PR 20-SEP-2000; 200OUS-0234052P.
PR 22-SEP-2000; 200OUS-0234567P.
PR 22-SEP-2000; 200OUS-0234923P.
PR 25-SEP-2000; 200OUS-0234924P.
PR 25-SEP-2000; 200OUS-0235077P.
PR 25-SEP-2000; 200OUS-0235082P.
PR 25-SEP-2000; 200OUS-0235134P.
PR 25-SEP-2000; 200OUS-0235280P.
PR 26-SEP-2000; 200OUS-0235377P.
PR 26-SEP-2000; 200OUS-0235538P.
PR 27-SEP-2000; 200OUS-0235711P.
PR 27-SEP-2000; 200OUS-0235720P.
PR 27-SEP-2000; 200OUS-0235840P.
PR 27-SEP-2000; 200OUS-0235863P.
PR 28-SEP-2000; 200OUS-0236028P.
PR 28-SEP-2000; 200OUS-0236030P.
PR 28-SEP-2000; 200OUS-0236033P.
PR 28-SEP-2000; 200OUS-0236049P.
PR 28-SEP-2000; 200OUS-0236109P.
PR 28-SEP-2000; 200OUS-0236111P.
PR 28-SEP-2000; 200OUS-0236842P.
PR 29-SEP-2000; 200OUS-0236891P.
PR 02-OCT-2000; 200OUS-0237172P.
PR 02-OCT-2000; 200OUS-0237173P.
PR 02-OCT-2000; 200OUS-0237278P.
PR 02-OCT-2000; 200OUS-0237294P.
PR 02-OCT-2000; 200OUS-0237295P.
PR 03-OCT-2000; 200OUS-0237316P.
PR 03-OCT-2000; 200OUS-0237425P.
PR 03-OCT-2000; 200OUS-0237598P.
PR 03-OCT-2000; 200OUS-0237604P.
PR 03-OCT-2000; 200OUS-0237606P.
PR 03-OCT-2000; 200OUS-0237608P.
PR 01-NOV-2000; 200OUS-0248667P.
PR 01-NOV-2000; 200OUS-0245084P.

XX

PS Claim 1; SEQ ID NO 646; 44pp; English.

XX

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
PI Soppet DR, Weaver Z;

XX

DR WPI; 2002-188264/24.

XX

PA (AVAL-) AVALON PHARM.

XX

PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.

XX

CC The present invention describes a method for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in AB161664
CC to ABI70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. MI can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of MI, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. MI can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour

XX

SQ Sequence 145831 BP; 37379 A; 35440 C; 35158 G; 37854 T; 0 U; 0 Other;

Query Match   11.3%; Score 98; DB 6; Length 145831;

Best Local Similarity   100.0%; Pred. No. 5.1e-27;
Matches   98; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
                ||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 10
ABT10149/c
ID   ABT10149 standard; cDNA; 145831 BP.
XX
AC   ABT10149;
XX
DT   04-DEC-2002 (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 283.
XX
KW   Human; breast specific gene; breast cancer; differential expression;
KW   cytostatic; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
XX
FN   WO200259271-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002NO-US002176.
XX
PR   25-JAN-2001; 2001US-0263757P.
PR   25-APR-2001; 2001US-0286090P.
PR   23-MAY-2001; 2001US-0292517P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Orr MS, Nation M, Diggans JC, Zeng W;
XX
DR   WPI; 2002-674803/72.
XX
PT   Diagnosing breast cancer in a patient comprises detecting the level of
PT   gene expression in cell or tissue samples, where a differential gene
PT   expression is indicative of breast cancer.
XX
CC   The present invention relates to methods of diagnosing breast cancer in a
CC   patient, which comprise detecting the level of expression in a tissue
CC   sample of two or more genes selected from those shown in ABT09867-
CC   ABT11112, where a differential expression of the genes indicates breast
CC   cancer. The methods are useful in diagnosing, treating, detecting the
CC   progression, and in monitoring treatment of breast cancer in patients.
CC   The methods are also useful as a screening tool for agents that modulate
CC   the onset or progression of breast cancer. The breast cancer genes may be
CC   used as diagnostic markers for the prediction or identification of the
CC   malignant state of breast tissue, for confirming the type and progression
CC   of cancer, and for drug screening and assays. The present sequence is a
CC   coding sequence of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 145831 BP; 37379 A; 35440 C; 35158 G; 37854 T; 0 U; 0 Other;

Query Match   11.3%; Score 98; DB 6; Length 145831;
Best Local Similarity   100.0%; Pred. No. 5.1e-27;
Matches   98; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   38600  AGTAGAGAATGGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy       87  CGCCCGCCTTCGGCCTTCCAAAGTGCTGGGATTACAGGC 124
             |||||||||||||||||||||||||||||||||||||||
Db   38540  CGCCCGCCTTCGGCCTTCCAAAGTGCTGGGATTACAGGC 38503

RESULT 11
AAK82012
ID    AAK82012    standard; DNA; 49561 BP.
XX
AC    AAK82012;
XX
DT    07-NOV-2001  (first entry)
XX
DE    Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36824.
XX
KW    Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW    cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS    Homo sapiens.
XX
PN    WO200157182-A2.
XX
PD    09-AUG-2001.
XX
PF    17-JAN-2001; 2001NO-US001354.
XX
PR    31-JAN-2000;  2000US-0179065P.
PR    04-FEB-2000;  2000US-0180628P.
PR    24-FEB-2000;  2000US-0184664P.
PR    02-MAR-2000;  2000US-0186350P.
PR    16-MAR-2000;  2000US-0189874P.
PR    17-MAR-2000;  2000US-0190076P.
PR    18-APR-2000;  2000US-0198123P.
PR    19-MAY-2000;  2000US-0205515P.
PR    07-JUN-2000;  2000US-0209467P.
PR    28-JUN-2000;  2000US-0214886P.
PR    30-JUN-2000;  2000US-0215135P.
PR    07-JUL-2000;  2000US-0216647P.
PR    07-JUL-2000;  2000US-0216880P.
PR    11-JUL-2000;  2000US-0217487P.
PR    11-JUL-2000;  2000US-0217496P.
PR    14-JUL-2000;  2000US-0218290P.
PR    26-JUL-2000;  2000US-0209963P.
PR    26-JUL-2000;  2000US-0220964P.
PR    14-AUG-2000;  2000US-0224518P.
PR    14-AUG-2000;  2000US-0244519P.
PR    14-AUG-2000;  2000US-0225213P.
PR    14-AUG-2000;  2000US-0225214P.
PR    14-AUG-2000;  2000US-0225266P.
PR    14-AUG-2000;  2000US-0225267P.
PR    14-AUG-2000;  2000US-0225270P.
PR    14-AUG-2000;  2000US-0225447P.
PR    14-AUG-2000;  2000US-0225757P.
PR    14-AUG-2000;  2000US-0225758P.
PR    14-AUG-2000;  2000US-0225759P.
PR    18-AUG-2000;  2000US-0226279P.
PR    22-AUG-2000;  2000US-0226279P.
PR    22-AUG-2000;  2000US-0226681P.
PR    22-AUG-2000;  2000US-0226868P.
PR    23-AUG-2000;  2000US-0227182P.
PR    30-AUG-2000;  2000US-0227709P.
PR    01-SEP-2000;  2000US-0228924P.
PR    01-SEP-2000;  2000US-0229287P.
PR    01-SEP-2000;  2000US-0229343P.
PR    01-SEP-2000;  2000US-0229344P.
PR    01-SEP-2000;  2000US-0229345P.
PR    05-SEP-2000;  2000US-0229509P.
PR    05-SEP-2000;  2000US-0229513P.
PR    06-SEP-2000;  2000US-0230437P.
PR    06-SEP-2000;  2000US-0230438P.
PR    08-SEP-2000;  2000US-0231242P.
```

```
PR    08-SEP-2000;  2000US-0231243P.
PR    08-SEP-2000;  2000US-0231244P.
PR    08-SEP-2000;  2000US-0231413P.
PR    08-SEP-2000;  2000US-0231414P.
PR    08-SEP-2000;  2000US-0232080P.
PR    08-SEP-2000;  2000US-0232081P.
PR    12-SEP-2000;  2000US-0231968P.
PR    14-SEP-2000;  2000US-0231969P.
PR    14-SEP-2000;  2000US-0232397P.
PR    14-SEP-2000;  2000US-0232398P.
PR    14-SEP-2000;  2000US-0232399P.
PR    14-SEP-2000;  2000US-0232400P.
PR    14-SEP-2000;  2000US-0232401P.
PR    14-SEP-2000;  2000US-0233064P.
PR    14-SEP-2000;  2000US-0233065P.
PR    21-SEP-2000;  2000US-0234223P.
PR    21-SEP-2000;  2000US-0234274P.
PR    25-SEP-2000;  2000US-0234997P.
PR    25-SEP-2000;  2000US-0234998P.
PR    26-SEP-2000;  2000US-0254848P.
PR    27-SEP-2000;  2000US-0235834P.
PR    27-SEP-2000;  2000US-0235836P.
PR    29-SEP-2000;  2000US-0236327P.
PR    29-SEP-2000;  2000US-0236367P.
PR    29-SEP-2000;  2000US-0236368P.
PR    29-SEP-2000;  2000US-0236369P.
PR    02-OCT-2000;  2000US-0236802P.
PR    02-OCT-2000;  2000US-0236370P.
PR    02-OCT-2000;  2000US-0237037P.
PR    02-OCT-2000;  2000US-0237038P.
PR    02-OCT-2000;  2000US-0237039P.
PR    03-OCT-2000;  2000US-0237040P.
PR    13-OCT-2000;  2000US-0239937P.
PR    20-OCT-2000;  2000US-0239938P.
PR    20-OCT-2000;  2000US-0240960P.
PR    20-OCT-2000;  2000US-0241221P.
PR    20-OCT-2000;  2000US-0241785P.
PR    20-OCT-2000;  2000US-0241786P.
PR    20-OCT-2000;  2000US-0241787P.
PR    20-OCT-2000;  2000US-0241808P.
PR    20-OCT-2000;  2000US-0241809P.
PR    01-NOV-2000;  2000US-0241826P.
PR    08-NOV-2000;  2000US-0246617P.
PR    08-NOV-2000;  2000US-0246474P.
PR    08-NOV-2000;  2000US-0246475P.
PR    08-NOV-2000;  2000US-0246476P.
PR    08-NOV-2000;  2000US-0246477P.
PR    08-NOV-2000;  2000US-0246478P.
PR    08-NOV-2000;  2000US-0246523P.
PR    08-NOV-2000;  2000US-0246524P.
PR    08-NOV-2000;  2000US-0246525P.
PR    08-NOV-2000;  2000US-0246526P.
PR    08-NOV-2000;  2000US-0246527P.
PR    08-NOV-2000;  2000US-0246528P.
PR    08-NOV-2000;  2000US-0246609P.
PR    08-NOV-2000;  2000US-0246610P.
PR    08-NOV-2000;  2000US-0246611P.
PR    17-NOV-2000;  2000US-0246613P.
PR    17-NOV-2000;  2000US-0249207P.
PR    17-NOV-2000;  2000US-0249208P.
PR    17-NOV-2000;  2000US-0249209P.
PR    17-NOV-2000;  2000US-0249210P.
PR    17-NOV-2000;  2000US-0249211P.
PR    17-NOV-2000;  2000US-0249212P.
PR    17-NOV-2000;  2000US-0249213P.
PR    17-NOV-2000;  2000US-0249214P.
PR    17-NOV-2000;  2000US-0249215P.
PR    17-NOV-2000;  2000US-0249216P.
PR    17-NOV-2000;  2000US-0249217P.
PR    17-NOV-2000;  2000US-0249218P.
PR    17-NOV-2000;  2000US-0249244P.
PR    17-NOV-2000;  2000US-0249245P.
```

Left column

```
PR   17-NOV-2000; 2000US-0242264P.
PR   17-NOV-2000; 2000US-0242665P.
PR   17-NOV-2000; 2000US-0242297P.
PR   17-NOV-2000; 2000US-0242299P.
PR   17-NOV-2000; 2000US-0242300P.
PR   01-DEC-2000; 2000US-0251602P.
PR   01-DEC-2000; 2000US-0251602P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251330P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0219888P.
PR   05-DEC-2000; 2000US-0256719P.
PR   05-DEC-2000; 2000US-0254179P.
PR   06-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0240977P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS   Disclosure; SEQ ID NO 36824; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting the
CC   nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention
XX
SQ   Sequence 49561 BP; 14811 A; 9800 C; 9812 G; 15138 T; 0 U; 0 Other;

Query Match             9.9%;  Score 86;  DB 4;  Length 49561;
Best Local Similarity   100.0%;
Matches   86;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       36 GGGGTTTCACCGTGTTAGCCAGAGATGGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCCT 95
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    13127 GGGGTTTCACCGTGTTAGCCAGAGATGGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCCT 13186

Qy       96 CGGCCTCCCAAAGTGCTGGGATTACA 121
            ||||||||||||||||||||||||||
Db    13187 CGGCCTCCCAAAGTGCTGGGATTACA 13212

RESULT 12
AAI62868
ID   AAI62868 standard; DNA; 267 BP.
XX
AC   AAI62868;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human genomic DNA SEQ ID NO 196.
XX
```

Right column

```
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
XX
PN   WO200155449-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001346.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUL-2000; 2000US-0216880P.
PR   14-JUL-2000; 2000US-0218920P.
PR   01-SEP-2000; 2000US-0225447P.
PR   06-SEP-2000; 2000US-0223443P.
PR   08-SEP-2000; 2000US-0204377P.
PR   08-SEP-2000; 2000US-0231243P.
PR   25-SEP-2000; 2000US-0234997P.
PR   29-SEP-2000; 2000US-0236367P.
PR   13-OCT-2000; 2000US-0239937P.
PR   08-NOV-2000; 2000US-0464767P.
PR   08-NOV-2000; 2000US-0464777P.
PR   08-NOV-2000; 2000US-0465525P.
PR   08-NOV-2000; 2000US-0465526P.
PR   08-NOV-2000; 2000US-0465528P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249265P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250301P.
PR   05-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0254179P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0240977P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-476225/51.
XX
PT   Novel plasma membrane associated proteins useful for diagnosing,
PT   treating, preventing and/or prognosing disorders related to the proteins,
PT   including cancer, immune response and neuronal disorders.
XX
PS   Example 2; SEQ ID NO 196; 532pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (AAI62752-AAI62961) and proteins
CC   (AAM42347-AAM42415) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases such as cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
```

CC   infections. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 267 BP; 42 A; 93 C; 72 G; 60 T; 0 U; 0 Other;

Qy   Query Match
Best Local Similarity 100.0%; Score 75; DB 4; Length 267;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   70 CTCCTGACCTTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193 CTCCTGACCTTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCATGAG 252

Qy   130 CCACTGCGCCCAGCC 144
          |||||||||||||||
Db   253 CCACTGCGCCCAGCC 267

RESULT 13
ID   AAH09234
AC   AAH09234;
XX
DT   26-JUN-2001 (first entry)
XX
DE   Human cDNA clone (3'-primer) SEQ ID NO:6069.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999; 99JP-00248036.
PR   27-AUG-1999; 99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   The present invention describes primer sets for synthesizing 5602 full-
PT   length cDNAs defined in the specification. Where a primer set comprises:
PT   (a) an oligo-dT primer and an oligonucleotide complementary to the
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 3; SEQ ID NO 6069; 2537pp + Sequence Listing; English.
XX

CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides.

CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; and AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 579 BP; 139 A; 141 C; 114 G; 180 T; 0 U; 5 Other;

Qy   Query Match
Best Local Similarity 100.0%; Score 74; DB 4; Length 579;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   60 TGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTA 133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 TGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTA 267

Qy   120 CAGGCATGAGCCAC 133
          ||||||||||||||
Db   268 CAGGCATGAGCCAC 281

RESULT 14
ID   AAH15304/C
AC   AAH15304;
XX
DT   26-JUN-2001 (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:13457.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999; 99JP-00248036.
PR   27-AUG-1999; 99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   The present invention describes primer sets for synthesizing 5602 full-
PT   length cDNAs defined in the specification. Where a primer set comprises:
PT   (a) an oligo-dT primer and an oligonucleotide complementary to the
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 13457; 2537pp + Sequence Listing; English.
XX

CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a

CC　polynucleotide which comprises a 3'-end sequence, where the
CC　oligonucleotide comprises at least 15 nucleotides and the combination of
CC　the 5'-end sequence/3'-end sequence is selected from those defined in the
CC　specification. The primer sets can be used in antisense therapy and in
CC　gene therapy. The primers are useful for synthesising polynucleotides,
CC　particularly full-length cDNAs. The primers are also useful for the
CC　detection and/or diagnosis of the abnormality of the proteins encoded by
CC　the full-length cDNAs.The primers allow obtaining of the full-length
CC　cDNAs easily without any specialised methods. AAH03166 to AAH3628 and
CC　AAH13633 to AAH18742 represent human cDNA sequences; and AAB95893
CC　represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC　oligonucleotides, all of which are used in the exemplification of the
CC　present invention
XX
SQ　Sequence 1555 BP; 488 A; 296 C; 332 G; 439 T; 0 U; 0 Other;

Query Match　　8.5%;　Score 74;　DB 4;　Length 1555;
Best Local Similarity　100.0%;　　Pred. No. 7.7e-18;
Matches　74;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　60　TGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTA 119
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　1348 TGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTA 1289

Qy　　120　CAGGCATGAGCCAC 133
　　　　　||||||||||||||
Db　　1288 CAGGCATGAGCCAC 1275

RESULT 15
ABK83562/c
ID　ABK83562 standard; cDNA; 139904 BP.
XX
AC　ABK83562;
XX
DT　14-AUG-2002　(first entry)
XX
DE　Human cDNA differentially expressed in granulocytic cells #133.
XX
KW　Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW　viral infection; parasitic infection; protozoal infection;
KW　fungal infection; sterile inflammatory disease; psoriasis;
KW　rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW　cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW　adult respiratory distress syndrome; inflammatory bowel disease;
KW　Crohn's disease; ulcerative colitis; periodontal disease;
KW　granulocyte activation; chronic inflammation; allergy.
XX
OS　Homo sapiens.
XX
PN　WO20228999-A2.
XX
PD　11-APR-2002.
XX
DR　WPI; 2002-435328/46.
XX
PF　03-OCT-2001; 2001WO-US030821.
XX
PR　03-OCT-2000; 2000US-0237189P.
XX
PA　(GENE-) GENE LOGIC INC.
XX
PI　Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
PS　Claim 1; SEQ ID NO 133; 114pp; English.
XX
CC　The invention relates to detecting differential (M1) granulocyte (GC) activation
CC　(GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC　DNA chip analysis as given in the specification, and comparing the

CC　expression level to an expression level in an unactivated GC, where the
CC　differential expression of Gs is indicative of GCA. Also included are
CC　modulating (M2) GA by contacting GC with an agent that alters the
CC　expression of at least one gene in Gs; (2) screening (M3) for an agent
CC　capable of modulating GCA or an inflammation (especially chronic) in a
CC　tissue, an allergic response in a subject, exposure of a subject to a
CC　pathogen or sterile inflammatory disease using the gene expression
CC　profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC　tissue, an allergic response in a subject, exposure of a subject to a
CC　pathogen or sterile inflammatory disease, by detecting the level of
CC　expression in a sample of the tissue of gene(s) from Gs, where the level
CC　of expression of the gene is indicative of inflammation; (4) treating
CC　(M5) an inflammation (especially chronic) or in a tissue, an allergic
CC　response in a subject, exposure of a subject to a pathogen or sterile
CC　inflammatory disease, by contacting a tissue having inflammation with an
CC　agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC　is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC　for screening an agent capable of modulating GCA preferably in an
CC　inflammation in a tissue; M4 is useful for detecting an inflammation
CC　(especially chronic) in a tissue, an allergic response in a subject,
CC　exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC　psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC　cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC　respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC　disease, ulcerative colitis, periodontal disease; also bacterial
CC　infection, viral infection, parasitic infection, protozoal infection,
CC　fungal infection and M5 is useful for treating one of the above
CC　conditions. The present sequence represents a gene differentially
CC　expressed in granulocytes. Note:The sequence data for this patent did
CC　not form part of the printed specification, but was obtained in
CC　electronic format directly from WIPO at
CC　ftp.wipo.int/pub/published_pct_sequences
XX
SQ　Sequence 139904 BP; 39268 A; 29759 C; 30173 G; 40704 T; 0 U; 0 Other;

Query Match　　8.5%;　Score 74;　DB 6;　Length 139904;
Best Local Similarity　100.0%;　　Pred. No. 3.1e-18;
Matches　74;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

Qy　　60　TGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTA 119
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　26820 TGGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCCCAAAGTGCTGGGATTA 26761

Qy　　120　CAGGCATGAGCCAC 133
　　　　　||||||||||||||
Db　　26760 CAGGCATGAGCCAC 26747

Search completed: May 24, 2004, 16:57:18
Job time : 337 secs

OM nucleic - nucleic search, using sw model

```
                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:      May 24, 2004, 16:51:33 ; Search time 360 Seconds
                                       [without alignments]
                                       10945.219 Million cell updates/sec

Title:            US-10-621-363-12
Perfect score:    867
Sequence:         1 tcgagtttttttttttttt............accaactcgtgccgaattc 867

Scoring table:    OLIGO_NUC
                  Gapop 60.0 , Gapext 60.0

Searched:         2953838 seqs, 2272363821 residues
Word size :       0
Total number of hits satisfying chosen parameters:     5907676
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Listing first 45 summaries

Database :    Published Applications NA:*
```

```
 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	867	100.0	867	10	US-09-774-639-61	Sequence 61, Appl
2	867	100.0	867	10	US-09-969-730-12	Sequence 12, Appl
3	867	100.0	867	17	US-10-621-363-12	Sequence 12, Appl
4	98	11.3	145831	9	US-09-969-708-79	Sequence 79, Appl
5	98	11.3	145831	9	US-09-954-456-2116	Sequence 2116, Ap
6	98	11.3	145831	10	US-09-873-367C-646	Sequence 646, Ap
7	98	11.3	145831	13	US-10-240-425-363	Sequence 363, App
8	75	8.7	267	9	US-09-860-670-196	Sequence 196, App
9	75	8.7	267	16	US-10-227-646-196	Sequence 196, App
10	71	8.2	269	9	US-09-867-701-10441	Sequence 10441, A
11	70	8.1	548	13	US-10-027-632-249911	Sequence 249911,
12	70	8.1	548	16	US-10-027-632-249911	Sequence 249911,
13	70	8.1	54200	13	US-10-087-192-190	Sequence 190, App
14	70	8.1	65454	17	US-10-293-864-11	Sequence 11, Appl
c 15	68	7.8	748	13	US-10-027-632-13366	Sequence 13366, A
c 16	68	7.8	748	13	US-10-027-632-13366	Sequence 13366, A
c 17	66	7.6	504	13	US-10-027-632-324709	Sequence 324709,
c 18	66	7.6	504	13	US-10-027-632-324710	Sequence 324710,
c 19	66	7.6	504	13	US-10-027-632-324711	Sequence 324711,
c 20	66	7.6	504	16	US-10-027-632-324709	Sequence 324709,
c 21	66	7.6	504	16	US-10-027-632-324710	Sequence 324710,
c 22	66	7.6	504	16	US-10-027-632-324711	Sequence 324711,
c 23	66	7.6	26657	9	US-09-810-673A-3	Sequence 3, Appli
c 24	66	7.6	26657	15	US-10-395-242-3	Sequence 3, Appli
c 25	66	7.6	148567	13	US-09-801-876B-3	Sequence 3, Appli
c 26	66	7.6	148567	13	US-10-667-442-3	Sequence 3, Appli
c 27	66	7.6	148567	15	US-10-254-869-3	Sequence 3, Appli
c 28	65	7.5	2984	13	US-10-108-260A-2336	Sequence 2336, Ap
c 29	65	7.5	3024	13	US-10-027-632-114298	Sequence 114298,
c 30	65	7.5	3024	13	US-10-027-632-114299	Sequence 114299,
c 31	65	7.5	3024	16	US-10-027-632-114298	Sequence 114298,
c 32	65	7.5	3024	16	US-10-027-632-114299	Sequence 114299,
c 33	65	7.5	8663	9	US-09-764-878-397	Sequence 397, App
c 34	65	7.5	8663	15	US-10-079-854-397	Sequence 397, App
c 35	65	7.5	60430	13	US-10-087-192-1732	Sequence 1732, Ap
c 36	64	7.4	201143	13	US-10-244-425-1099	Sequence 1099, Ap
c 37	64	7.4	264	10	US-09-764-891-8680	Sequence 8680, Ap
c 38	64	7.4	290	9	US-09-764-847-1453	Sequence 1453, Ap
c 39	64	7.4	290	9	US-09-764-847-1454	Sequence 1454, Ap
c 40	64	7.4	290	15	US-10-092-154-1453	Sequence 1453, Ap
c 41	64	7.4	290	15	US-10-092-154-1454	Sequence 1454, Ap
c 42	64	7.4	291	9	US-09-764-847-1455	Sequence 1455, Ap
c 43	64	7.4	291	9	US-09-764-847-1456	Sequence 1456, Ap
c 44	64	7.4	291	15	US-10-092-154-1455	Sequence 1455, Ap
c 45	64	7.4	291	15	US-10-092-154-1456	Sequence 1456, Ap

```
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

ALIGNMENTS

```
RESULT 1
US-09-774-639-61
; Sequence 61, Application US/09774639
; Publication No. US2003003555A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 90 Human Secreted Proteins
;  FILE REFERENCE: P201IP1
;  CURRENT APPLICATION NUMBER: US/09/774,639
;  CURRENT FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 371
;  SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 61
;  LENGTH: 867
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: SITE
;  LOCATION: (831)
;  OTHER INFORMATION: n equals a,t,g, or c
US-09-774-639-61

Query Match         100.0%;  Score 867;  DB 10;  Length 867;
Best Local Similarity   100.0%;
Matches   867;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAATGGGGTTTCACCGGTGTTAGCCAGGAT  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGAATGGGGTTTCACCGGTGTTAGCCAGGAT  60

Qy   61 GGTCTCGGAACTCCTGACCTCGTGATCCGCCCGCCTTCCGAAAGTGCTGGGATTAC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGTCTCGGAACTCCTGACCTCGTGATCCGCCCGCCTTCCGAAAGTGCTGGGATTAC 120

Qy  121 AGGCATGAGCCACTCGCGCCCAGCCGGGTCTTTTAAACATTCCCAGGACTGTGTACAGCCAA 180
```

RESULT 2
US-09-969-730-12
Sequence 12, Application US/0969730
Publication No. US20030054443A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 90 Human Secreted Proteins
FILE REFERENCE: PZ013P2
CURRENT APPLICATION NUMBER: US/09/969,730
CURRENT FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 09/774,639
PRIOR FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: 60/238,291
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 09/244,112
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: PCT/US98/16235
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/056,371
PRIOR FILING DATE: 1997-08-19
PRIOR APPLICATION NUMBER: 60/056,732

PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/056,366
PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/056,364
PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/056,370
PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/056,367
PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/056,365
PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/056,731
PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/056,557
PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/056,563
PRIOR FILING DATE: 1997-08-19 PRIOR APPLICATION NUMBER: 60/055,970
PRIOR FILING DATE: 1997-08-18 PRIOR APPLICATION NUMBER: 60/055,986
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/055,311
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/054,808
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/054,803
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/054,804
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/054,809
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/054,806
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/054,798
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/055,310
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/055,309
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/055,312
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/054,807
PRIOR FILING DATE: 1997-08-05 PRIOR APPLICATION NUMBER: 60/055,386
PRIOR FILING DATE: 1997-08-05

SOFTWARE: Patentin Ver. 2.0
NUMBER OF SEQ ID NOS: 373

SEQ ID NO 12
LENGTH: 867
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (831)
OTHER INFORMATION: n equals a,t,g, or c
US-09-969-730-12

Query Match 100.0%; Score 867; DB 10; Length 867;
Best Local Similarity 100.0%;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[DNA sequence alignment blocks Qy/Db, positions 1–867]

```
;   PRIOR APPLICATION NUMBER: 60/056,366
;   PRIOR FILING DATE: 1997-08-19
;   PRIOR APPLICATION NUMBER: 60/056,364
;   PRIOR FILING DATE: 1997-08-19
;   PRIOR APPLICATION NUMBER: 60/056,370
;   PRIOR FILING DATE: 1997-08-19
;   Remaining Prior Application data removed - See File Wrapper or PALM.
;   NUMBER OF SEQ ID NOS: 373
;   SOFTWARE: Patentln Ver. 2.0
;
;   SEQ ID NO 12
;   LENGTH: 867
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (831)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-621-363-12

Query Match             100.0%;  Score 867;  DB 17;  Length 867;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  867;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGAGTTTTTTTTTTTTTTTTTTTTTTTAAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGAT 60

Qy         61 GGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCGCCCAAAGTGCTGGGATTAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGTCTCGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCGCCCAAAGTGCTGGGATTAC 120

Qy        121 AGGCATGAGCCACTGCGCCCAGCCGGCGTCTTTTAAACAATCCCCAGGACTGTACAGCCAA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCATGAGCCACTGCGCCCAGCCGGCGTCTTTTAAACAATCCCCAGGACTGTACAGCCAA 180

Qy        181 CCCATACTCACCTGACATTTGGGAACTCCCCCCACGGGCCATAACTGATCTGCAGAGGTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCATACTCACCTGACATTTGGGAACTCCCCCCACGGGCCATAACTGATCTGCAGAGGTA 240

Qy        241 AGACCAAGCAGAATGGGGGATTCACATCTAAGGTCTGGTGATGGGCTGAAGGAAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGACCAAGCAGAATGGGGGATTCACATCTAAGGTCTGGTGATGGGCTGAAGGAAG 300

Qy        301 AAGAATCAGCGAACAAAAGCCTCTAGGTCTTTTCTTACCACAAACACCTCTCTGCCCACCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGAATCAGCGAACAAAAGCCTCTAGGTCTTTTCTTACCACAAACACCTCTCTGCCCACCT 360

Qy        361 GCTTTGAAAGGGGCAGAAGTAGTAGTGGGCGGAGCTGCCCACCTGGCTACAGTGAAGGGATCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTTTGAAAGGGGCAGAAGTAGTAGTGGGCGGAGCTGCCCACCTGGCTACAGTGAAGGGATCT 420

Qy        421 GGAGAAATACTCACACTTTGAGGTGCTCGCCCCTCTTCATCAGCCAGTGAAGGGATCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAGAAATACTCACACTTTGAGGTGCTCGCCCCTCTTCATCAGCCAGTGAAGGGATCT 480

Qy        481 AATGACCCCACGGGAGCTTACACAAGTYCAAACAGGCCCAAATGCATTCATGAGGACAGGGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATGACCCCACGGGAGCTTACACAAGTYCAAACAGGCCCAAATGCATTCATGAGGACAGGGG 540

Qy        541 GAGGCCAAAGGACTCCGGAGGAGAGGGGCCCAATAAGGCTGGTGTCTATTTCCGATCCATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGCCAAAGGACTCCGGAGGAGAGGGGCCCAATAAGGCTGGTGTCTATTTCCGATCCATA 600

Qy        601 GAGAGAGCAGAGGTGGGCAGGCCCTTTGATTAATGTATCATTCTTGAATGCAAAGCTTCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAGAGCAGAGGTGGGCAGGCCCTTTGATTAATGTATCATTCTTGAATGCAAAGCTTCA 660

Qy        661 AAATCCCGGGTATGCCCGGGTGAGAATGAGCAGGAGCTAACACCTGGGGTGTCATGGCCAAGCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAATCCCGGGTATGCCCGGGTGAGAATGAGCAGGAGCTAACACCTGGGGTGTCATGGCCAAGCCT 720

Qy        721 CCAGGGCCGACTCGGCCAGAGACAGATCCGCAAGAGGACTAACACCTGGGGTGTCATGGCCAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCAGGGCCGACTCGGCCAGAGACAGATCCGCAAGAGGACTAACACCTGGGGTGTCATGGCCAA 780

Qy        781 GCCCACTCGGATTTGAACCCCGGGCTCCTCAAGGTCAGCTGTGTGTAGCCTTGANTGAAAYCACC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCCCACTCGGATTTGAACCCCGGGCTCCTCAAGGTCAGCTGTGTGTAGCCTTGANTGAAAYCACC 840

Qy        841 TGCTATGACCAATCTCGTGCCGAATTC 867
              |||||||||||||||||||||||||||
Db        841 TGCTATGACCAATCTCGTGCCGAATTC 867


RESULT 3
US-10-621-363-12

;   Sequence 12, Application US/1062l363
;   Publication No. US20040023283A1
;   GENERAL INFORMATION:
;   APPLICANT: Ruben et al.
;   TITLE OF INVENTION: 90 Human Secreted Proteins
;   FILE REFERENCE: PZ013P2C1
;   CURRENT APPLICATION NUMBER: US/10/621,363
;   CURRENT FILING DATE:     2003-07-18
;   PRIOR APPLICATION NUMBER: 09/969,730
;   PRIOR FILING DATE:     2001-10-06
;   PRIOR APPLICATION NUMBER: 09/774,639
;   PRIOR FILING DATE:     2001-02-01
;   PRIOR APPLICATION NUMBER: 60/238,291
;   PRIOR FILING DATE:     2000-10-06
;   PRIOR APPLICATION NUMBER: 09/244,112
;   PRIOR FILING DATE:     1999-02-04
;   PRIOR APPLICATION NUMBER: PCT/US98/16235
;   PRIOR FILING DATE:     1998-08-04
;   PRIOR APPLICATION NUMBER: 60/056,371
;   PRIOR FILING DATE:     1997-08-19
;   PRIOR APPLICATION NUMBER: 60/056,732
;   PRIOR FILING DATE:     1997-08-19
```

RESULT 4
US-09-969-708-79/c
; Sequence 79, Application US/0996970B
; Patent No. US2002010253ZA1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 79
; LENGTH: 145831
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-79

  Query Match   11.3%; Score 98; DB 9; Length 145831;
  Best Local Similarity   100.0%; Pred. No. 5.2e-40;
  Matches   98; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
               ||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 5
US-09-954-456-2116/c
; Sequence 2116, Application US/0995445G
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711

Db   721 CCAGGGGCCGACTGGCCAGAGACAGATCCGCAAGAGGCTTCTGCAGCCAGCTCTGGTGCCGTGCCAA 780

Qy   781 GCCACTCGGGATTTGAACCCCGCGCCTCCTCCAAAGTCAGCTGTGAGCCTTTGAANTGAAYGAAYCACC 840
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GCCACTCGGATTTGAACCCCGGCGCCTCCTCCAAAGTCAGCTGTGAGCCTTTGAANTGAAYCACC 840

Qy   841 TGCTATGACCAANTCTCGTGCCGAATTC 867
               ||||||||||||||||||||||||||||
Db   841 TGCTATGACCAANTCTCGTGCCGAATTC 867

RESULT 6
US-09-873-367C-646/c
; Sequence 646, Application US/0987336?C
; Publication No. US2003016583?A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 646
; LENGTH: 145831
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-646

  Query Match   11.3%; Score 98; DB 10; Length 145831;
  Best Local Similarity   100.0%; Pred. No. 5.2e-40;
  Matches   98; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
               ||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 38503

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2116
; LENGTH: 145831
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-2116

  Query Match   11.3%; Score 98; DB 9; Length 145831;
  Best Local Similarity   100.0%; Pred. No. 5.2e-40;
  Matches   98; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   27 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGAACTCCTGACCTCGTGATC 38541

Qy   87 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 124
               ||||||||||||||||||||||||||||||||||||||
Db   38540 CGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 38503

RESULT 7
US-10-240-425-363/c
; Sequence 363, Application US/10240425
; Publication No. US20040035502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 363
; LENGTH: 145831
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040035502A1 AF001548

Query Match   11.3%;   Score 98;   DB 13;   Length 145831;
Best Local Similarity 100.0%;   Pred. No. 5.2e-40;
Matches   98;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   27 AGTAGAGAGGGGGGTTTCACGTGTGTTAGCGAGGAGGAGGTCTCGAACTCCTCGACCTCCTGATCC 86
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38600 AGTAGAGAGGGGGGTTTCACGTGTGTTAGCGAGGAGGAGGTCTCGAACTCCTCGACCTCCTGATCC 38541

Qy   87 CGGCGGGGCCTCGGGCCTTCCCAAAGTGCTCGGGATTACACGGC 124
            ||||||||||||||||||||||||||||||||||||||||||||
Db   38540 CGGCGGGGCCTCGGGCCTTCCCAAAGTGCTCGGGATTACACGGC 38503

RESULT 8
US-09-860-670-196
; Sequence 196, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 196
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-196

Query Match   8.7%;   Score 75;   DB 9;   Length 267;
Best Local Similarity 100.0%;   Pred. No. 5.3e-28;
Matches   75;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   70 CTCCTGACCTCGTGATCCGGCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAG 129
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193 CTCCTGACCTCGTGATCCGGCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAG 252

Qy   130 CCACTGCGCCCAGCC 144
            |||||||||||||||
Db   253 CCACTGCGCCCAGCC 267

RESULT 9
US-10-227-646-196
; Sequence 196, Application US/10227646
; Publication No. US20030235829A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/10/227,646
; CURRENT FILING DATE: 2002-08-26
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US/09/860,670
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 60/205,515
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 196
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-227-646-196

Query Match   8.7%;   Score 75;   DB 16;   Length 267;
Best Local Similarity 100.0%;   Pred. No. 5.3e-28;
Matches   75;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   70 CTCCTGACCTCGTGATCCGGCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAG 129
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   193 CTCCTGACCTCGTGATCCGGCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAG 252

Qy   130 CCACTGCGCCCAGCC 144
            |||||||||||||||
Db   253 CCACTGCGCCCAGCC 267

RESULT 10
US-09-867-701-10441/c
; Sequence 10441, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10441
; LENGTH: 269
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-10441

Query Match   8.2%;   Score 71;   DB 9;   Length 269;
Best Local Similarity 100.0%;   Pred. No. 6e-26;
Matches   71;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   74 TGACCTCGTGATCCGGCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC 133
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   163 TGACCTCGTGATCCGGCCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC 104

Qy   134 TGCGCCCAGCC 144
            |||||||||||
Db   103 TGCGCCCAGCC 93

RESULT 11
US-10-027-632-249911/c

; Sequence 249911, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 249911
LENGTH: 548

; TYPE: DNA
; ORGANISM: Human
US-10-027-632-249911

Query Match   8.1%; Score 70;   DB 13; Length 548;
Best Local Similarity   100.0%; Pred. No. 1.9e-25;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   70 CTCCTGACCTCGTGATCCGGCCGCCCCCGCCTCGGCCTCCCAAAGTGCTGGGGATTACAGGCATGAG 129
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 548 CTCCTGACCTCGTGATCCGGCCGCCCCCGCCTCGGCCTCCCAAAGTGCTGGGGATTACAGGCATGAG 489

Qy 130 CCACTGCGCC 139
        ||||||||||
Db 488 CCACTGCGCC 479

RESULT 12
US-10-027-632-249911/c
; Sequence 249911, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30

; TYPE: DNA
; ORGANISM: Human
US-10-027-632-249911

Query Match   8.1%; Score 70;   DB 16; Length 548;
Best Local Similarity   100.0%; Pred. No. 1.9e-25;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   70 CTCCTGACCTCGTGATCCGGCCGCCCCCGCCTCGGCCTCCCAAAGTGCTGGGGATTACAGGCATGAG 129
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 548 CTCCTGACCTCGTGATCCGGCCGCCCCCGCCTCGGCCTCCCAAAGTGCTGGGGATTACAGGCATGAG 489

Qy 130 CCACTGCGCC 139
        ||||||||||
Db 488 CCACTGCGCC 479

RESULT 13
US-10-087-192-190
; Sequence 190, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
CURRENT APPLICATION NUMBER: US/10/087,192
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-190
; LENGTH: 54200

Query Match   8.1%; Score 70;   DB 13; Length 54200;
Best Local Similarity   100.0%; Pred. No. 1.4e-25;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   70 CTCCTGACCTCGTGATCCGCCCGCCCCCGCCTCGGCCTCCCAAAGTGCTGGGGATTACAGGCATGAG 129
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4549 CTCCTGACCTCGTGATCCGCCCGCCCCCGCCTCGGCCTCCCAAAGTGCTGGGGATTACAGGCATGAG 4608

Qy 130 CCACTGCGCC 139
        ||||||||||
Db 4609 CCACTGCGCC 4618

RESULT 14
US-10-293-864-11/c
; Sequence 11, Application US/10293864
; Publication No. US20040092465A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: MODULATION OF HUNTINGTIN INTERACTING PROTEIN 1 EXPRESSION
; FILE REFERENCE: RTS-0432
CURRENT APPLICATION NUMBER: US/10/293,864
CURRENT FILING DATE: 2002-11-11
; NUMBER OF SEQ ID NOS: 165
SEQ ID NO 11
; LENGTH: 65454
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
US-10-293-864-11

Query Match   8.1%; Score 70;   DB 17; Length 65454;

```
Best Local Similarity   100.0%;  Pred. No. 1.3e-25;
Matches  70;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       70 CTCCTGACCTCGTGATCCGGCCCGCCCGCCTCCGGCCTCCCAAAAGTGCTGGGGATTACAGGCATGAG 129
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     9836 CTCCTGACCTCGTGATCCGGCCCGCCCGCCTCCGGCCTCCCAAAAGTGCTGGGGATTACAGGCATGAG 9777

Qy      130 CCACTGCGCC 139
            ||||||||||
Db     9776 CCACTGCGCC 9767
```

RESULT 15
US-10-027-632-13366/c

; Sequence 13366, Application US/10027632
; Publication No. US20020183711A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: PatentIn for Windows Version 4.0
; SEQ ID NO 13366
; LENGTH: 748
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-13366

Query Match 7.8%; Score 68; DB 13; Length 748;
Best Local Similarity 100.0%; Pred. No. 2e-24;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   72 CCTGACCTCCGTGATCCGGCCCGCGCCCTCGGCGCCTCCCAAAAGTGCTGGGGATTACAGGCATGAGCC 131
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   356 CCTGACCTCGTGATCCGGCCCGCGCCCTCGGCGCCTCCCAAAAGTGCTGGGGATTACAGGCATGAGCC 297

Qy   132 ACTGCGCC 139
           ||||||||
Db   296 ACTGCGCC 289
```

Search completed: May 24, 2004, 18:26:31
Job time : 363 secs